```
2, 2002, 09:11:52; Search time 101.03 Seconds (without alignments) 1052.139 Million cell updates/sec
GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                 OM protein - protein search, using sw model
                                                                                                                                                                         Run on:
```

1 MDRSENIRLTMRSRRLVSLL.....RASFGSVNPATPTADTYLQX BLOSUM62 Gapop 10.0 , Gapext 0.5 US-09-917-376-1 5135 Perfect score: Scoring table: Sequence:

Total number of hits satisfying chosen parameters: 747574 seqs, 111073796 residues

Searched:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

1: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
110: /SIDSI/gcgdata/hold-geneseqy-embl/AA1981.DAT:*
121: /SIDSI/gcgdata/hold-geneseqy-embl/AA1981.DAT:*
122: /SIDSI/gcgdata/hold-geneseqy-embl/AA1991.DAT:*
123: /SIDSI/gcgdata/hold-geneseqy-embl/AA1991.DAT:*
143: /SIDSI/gcgdata/hold-geneseqy-embl/AA1991.DAT:*
144: /SIDSI/gcgdata/hold-geneseqy-embl/AA1991.DAT:*
145: /SIDSI/gcgdata/hold-geneseqy-embl/AA1991.DAT:*
146: /SIDSI/gcgdata/hold-geneseqy-embl/AA1991.DAT:*
147: /SIDSI/gcgdata/hold-geneseqy-embl/AA1991.DAT:*
148: /SIDSI/gcgdata/hold-geneseqy-embl/AA1991.DAT:*
149: /SIDSI/gcgdata/hold-geneseqy-embl/AA1991.DAT:*
149: /SIDSI/gcgdata/hold-geneseqy-embl/AA1999.DAT:*
140: /SIDSI/gcgdata/hold-geneseqy-embl/AA1999.DAT:*
141: /SIDSI/gcgdata/hold-geneseqy-embl/AA1999.DAT:*
142: /SIDSI/gcgdata/hold-geneseqy-embl/AA1999.DAT:*
143: /SIDSI/gcgdata/hold-geneseqy-embl/AA1999.DAT:*
144: /SIDSI/gcgdata/hold-geneseqy-embl/AA1999.DAT:*
145: /SIDSI/gcgdata/hold-geneseqy-embl/AA1999.DAT:*
146: /SIDSI/gcgdata/hold-geneseqy-embl/AA1999.DAT:*
147: /SIDSI/gcgdata/hold-geneseqy-embl/AA1999.DAT:*
148: /SIDSI/gcgdata/hold-geneseqy-geneseqp-embl/AA1999.DAT:*
159: /SIDSI/gcgdata/hold-geneseqy-geneseqp-embl/AA1999.DAT:*
150: /SIDSI/gcgdata/hold-geneseqy-geneseqp-embl/AA1999.DAT:*
151: /SIDSI/gcgdata/hold-geneseqy-geneseqp-embl/AA1999.DAT:*
151: /SIDSI/gcgdata/hold-geneseqy-geneseqp-embl/AA1999.DAT:*
151: /SIDSI/gcgdata/hold-geneseqy-geneseqp-embl/AA1999.DAT:*
151: /SIDSI/gcgdata/hold-geneseqy-geneseqp-embl/AA1999.DAT:* A_Geneseq_032802:* Database

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Novel human diagno Scaffoldin protein C. thermocellum ce C. thermocellum ol Drosophila melanog Propionibacterium Cellulase AE-1. A Glutamic acid rich Truncated cellulas Truncated cellulas Corrected Bacillus Description SUMMARIES AAW15238 AAW43108 AAW43106 ABB71513 AAU44346 AAW18790 AAY13493 AAR15625 AAY13492 ABG27250 AAB72314 ü Match Length DB Query Score Š. Result

Streptococcus pneu	u	Leishmania major P	Drosophila melanog	Staphylococcus aur	Acidothermus cellu	Novel human diagno	ORFX	Acidothermus cellu	A. cellulolyticus	Acidothermus cellu	Acidothermus cellu	S. aureus ClfB pro	Streptococcus pneu		Pectate lyase CBD	Amino acid sequenc	Bankia gouldi endo	Drosophila melanog	Propionibacterium	Drosophila melanog	NK-1 cellulase. B	Human MUC11 polype	C900P predicted am	A mannanase-linker	Protein encoded by	Novel endoglucanas	Streptococcus pneu	Drosophila melanog	Saccharothrix aust	. Truncated cellulas	Arabidopsis thalia	Arabidopsis thalia	
AAY81609	AA002124	AAB74609	ABB70775	AAU37120	AAB48788	ABG14000	AAB40945	AAB48787	AAR89927	AAB48786	AAY69508	AAY08640	AAW14574	AAY28850	AAY43218	AAW59912	AAW34987	ABB65262	AAU51992	ABB65175	AAR42122	AAY59288	AAM24513	AAY54123	AAB74209	AAR13227	AAW14570	ABB66878	AAW95602	AAY13494	AAG08664	AAG46507	10CCBUKK
21	22	22	22	22	22	22	21	22	17	22	21	20	18	50	20	19	19	22	22	22	14	21	22	21	22	12	18	22	20	20	21	21	5
1237	88	237	1357	2344	521	298	1532	521	521	521	562	918	206	493	493	1291	875	1805	311	1180	499	957	957	476	731	200	183	2586	490	616	317	333	0 7 7
4.3	4.2	4.2	4.1	4.1	4.1	4.1	4.1	4.1	4.1	4.1	4.1	4.1	4.1	4.1	4.1	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	3.9	3.9	•	3.9	3.9	•	•	3.9	•	c
219	218	216	213	213	212.5	212.5	212	211.5	210.5	210.5	210.5	209	208.5	208.5	208.5	207.5	206.5	206	205	205	204.5	204.5	204.5	202.5	202	200.5	200	200	99.		199	6	С
12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	٧.

 ζ Truncated ζ ellulases comprising amino acid sequence. AAY13493 standard; Protein; 1751 AA 30-JUL-1999 (first entry) AAY13493; AAY13493

ALIGNMENTS

RESULT

Cellulase; proteinase; truncated; Cel B5; Cel B4/5; Cel E1; Cel E1/2; Cel 1/2/3; Cel 6; Cel E3/B5; laundry detergent; stain-removing; cotton-containing fabric; stonewashing. 98EP-0810919 Unidentified 15-SEP-1998; EP921188-A2 09-JUN-1999

, Bergquist PL, Daniels RM, Farrington GK; Morgan H, Williams DP; (CLRN) CLARIANT FINANCE BVI LTD. WPI; 1999-315403/27 N-PSDB; AAX55662. Anderson P, Gibbs MD,

97US-0932571

19-SEP-1997;

New truncated cellulase proteins, useful in detergents and for producing 'stonewashed' denim

~

```
ADGSRF---VW--APGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 2; Fig 3; 8pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                90JP-0045465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   90JP-0045465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Aeromonas strain no. 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    934 CGNIRASFGSVNPA 947
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----svassinpa 947
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (OJIP ) OJI PAPER KK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1991-373412/51.
N-PSDB; AAQ15178.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                782 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Cellulase AE-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JP03251174-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          38 - NOV - 1991.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     38-FEB-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-FEB-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        17-MAR-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR15625
                                                                                                                                                                                                                                                                                                                                             814
                                        607
                                                                                   635
                                                                                                                              753
                                                                                                                                                                                                                                                                                                  825
                                                                                                                                                                                                                                                                                                                                                                                          842
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
AAR15625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                    g
                                                                                   δ
                                                                                                                            g
                                                                                                                                                                      Ω
                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                          ογ
                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                   The invention relates to a recombinant cellulase active protein free of proteinases of native thermophilic and alkalinophilic origin, comprising the truncated sequences Cell B5, Cell B45, Cell B1/5. Cell B1/3.

The truncated sequences Cell B5, Cell B5, Cell B1/5.

Cell 6 or Cell E3/B5, or a stability region from one of the defined full-cell b1/8.

Cell 6 or Cell E3/B5, or a stability region from one of the defined full-cell B1/8.

Cell 6 or Cell E3/B5, or a stability region from one of the defined from amino acid W135 in the sequence shown in AAY13492; Cell E1 extends from amino acid W133 to E1/2.

CEL/2.3 extends from amino acid W133 to E1/2 extends from manio acid W133 to E1/2.

CEL/2.3 extends from wanio acid W133 to E1/2 extends from manio acid W133 to E1/2.

CEL/2.3 extends from wanio acid W133 to E1/2 extends from manio acid W133 to E1/2.

CEL/2.3 extends from wanio acid W133 to E1/2 extends from manio acid W133 to E1/2.

CEL/2.3 extends from wanio acid W133 to E1/2 extends from manio acid W133 to E1/2.

CEL/2.3 extends from manio acid W133 to E1/2 extends from manio acid W133 to E1/2 extends from wanio acid W133 to E1/2 extends from acid W1340. The new consuming acid E4/B5 is shown in AAY13494. The new materials including cotton-containing fabrics. They are especially useful for processing of textiles where callulose breakdown is required. The new truncated enzymes show reduced redeposition of dye compared to using contruncated cellulase compositions.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                186 svvaet----aaslavasvvikernsqkaasylqhakdlfefadttrsdagytaatgfy 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        241 tsggfiddlgw---aavwly1atndssyltkaeelmseyangtntwtqcwddvrygtlim 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       298 lakitgkel--ykgavernldhwtdritytpkgmayltgw----gslryattaaflacvy 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --ITPLPFKLGGNMPGR----GMGERLAVDPNNDNILYFGAPSG 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIAT-- 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTAN.DYF-----GYSGLTIDRQ 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QPNPPVPSPKLGW----MDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPM 469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----wfk-----1vngg--sss 708
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---TTQP---YTWSNVAI 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GGGGFVDGIVFNEGAPGILYVRTD------IGGMYRW-----DAANGRWIP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        195 KGLWRSTDSGATWSQMTNFPD-----VGTYIANPIDITGYQSDIQGVVW--VAFDKSSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      453 ivgalak------ronyqlyggepid--dfkaietptndeifveskf
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVL-DISAEPWLTFGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -pdvvkvdt-----yyleggkisgpyvwd----knrniyyvlvdfsgtk----i
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     576 ypggevehkkqaqfkisvpqgypwdptn-dpsykgltsqleknkyiaaydnnn----1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VKGLEETAVNDLISP-----PSGAPLISALGDLGGFTHADVTAVPSTIFT---SPVFT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           vwglepgaatstpaptstptptptptvta------tptptptptptgsp--g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LLDWVGWNNWGYNGVVSIAADPINTNKVWAAVGM-YTNSWDPNDGAILRSSDQG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           352 adwsgcdsnkktkylnfaksgidyalgst--grsfvvgfgtnypghphkn-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 6.5%; Score 334; DB 20; Length 1751; Best Local Similarity 21.2%; Pred. No. 2e-09; Matches 219; Conservative 125; Mismatches 314; Indels 376;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SLLAATASFAVAAALGVLPIAITASPAHAA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              tgsgvkvlyknnetsastgsirp-----
                  Claim 7; Page 37-41; 65pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1751 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    153 ATWQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              675
```

ò g à g ò g δ g δy g δ Dp g

ò

g

οy

g

δ

g

δy

δ

```
753
                                                                                                                                                                          v----sgadyy-levgf-ssgagglg----pgkdtgdigvrfnkndwsnyngaddwsw 800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAIG 933
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The sequence was deduced from the gene which was sequenced from plasmid, pAEC 1, prepd. by ligating chromosomal DNA contg. the gene (obtd. from Aeromonas) into pUCL8. The protein has amol. wt. of 81,000 (SDS-PAES) and an optimum pH near to 5 when carboxymethylcellulose is the substrate. The N-terminal sequence: GIHADT-has been confirmed by Edman degradation. The gene can be used to produce recombinant enzyme which is used for the effective utilisation of biomass resources and the mfr. of pharmaceuticals and foodstuffs, and also for the detergent and deinking of waste paper.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q-----tpaptstatpt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    vdlsrvkirywytvdgdkpqsavc----dw-----aqigasnvt-fnfvklssg
                                                                                            TFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKE-GDLWLAASSGLYHSTNGGSSWSA
                                                                                                                                                                                                                                                                  694 ITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWG
                                                                                                                                                                                                                                                                                                                                                         801 lgsmtn----ygenakvtly-----vdg------vlv----wg
                                                                                                                                                                                                                                                                                                                                                                                                                                            754 QAITGDHANLRRVYIGTNGRGIVYGDIGGAPSGSPSVSPSVSPSASPSLSPSPSPSPSPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Cellulase AE-1 for e.g. mfr. of pharmaceuticals and foodstuffs of opt. pH when carboxy:methyl:cellulose is used as substrate.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Detergents; pharmaceuticals; deinking; carboxymethylcellulose
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR15625 standard; Protein; 782 AA.
```

us-09-917-376-1.rag

16-MAY-2001 (first entry)

XX AC AC AC XX	XX		90 X	NA X d	XX PF	E 6	XX A G	XX LI G	AQ XX	Id d															QY	qa	Qy	
Ouery Match 6.3%; Score 323; DB 12; Length 782; Best Local Similarity 19.7%; Pred. No. 3.1e-09; Matches 182; Conservative 106; Mismatches 294; Indels 344; Gaps	121 PINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWOITPLPFFKLGGNMPGRGMGERLAVD 180 1	PUNDNILYFGAPSGRGLWRSTDSGATWSQWINFPDV	qasadrihhadfssvteggtyrlwvsglgesvpfaingnlypdlpseameyfyfhrmgvd	21/GITIANPIDITGIQDDIQGYWWAPURSSSBLGGARXIIFYGYADPNNPYWRSKBGG- 2/3 	ATWQAVPGA-PTGFIPHKGVFDPVNHVLXIATSNTG-GPYD	giypvnqaisawtlvnlyerypgafpdgslnipesgngvpdildevlfgstfmkgvmpst	313 GSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWP 369 :	DTI-IFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQFNPPVPS	311 rlikpykeaeaaemwgiakeawarasanpnvlytsqtpdapg 352	ΡK	353 ggdyddirtsddryaaaaelylt375	482 ISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFD 541	GKNWFQGSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVG	; ; ; ;	02 FGNSWAASQGVPANAQIRSDRVNPKTFYALS-NGTFYRSTDGGVTFQPVAAGLPSSGAVG	459 kgydwgfmgnnamillgysydlsqnldylkamnrgmdyfmgnnamr 506	661 VMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVV 720 :::	CHICCUMCAVDENDOCONOCOMMUNT INDIDUCES	gricovication of the control of the	Ö		811 SPSPSSSPSSSPSSPSPSPSPSPSRSPSPSASPSPSSSSPSPSSSPSPTPSSSPVS 870	dudədədi	GGVKVQXKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRVWFTRD	åď	931 AIGCGNIRASFGSVNPATPTADTYLQ 956	snlngafvkmnpgkan	
Quer Best Matc	oy d	, y	a :	S S	δο i	ag .	Oy Dp	ογ	g	Οy	qq	Oy Op	λ̈́O	qq	δŏ	qq	oy B	3 8	g 8	οy	QQ	Qy	qq	δλ	qq	Qy	QQ	

AAB72314 standard; Protein; 1749 AA.

RESULT AAB72314

```
This invention relates to defence-related signalling genes isolated from the sunflower (Helianthus annuus). The genes encode a neoxanthin cleavage cargyme (MCE), an antho acid permease (AAP) and a glutemic acid rich protein (GRP). The signalling gene is useful for increasing the protein (GRP). The signalling gene is useful for increasing the casistance of a plant to a pathogen such as fungus, virus, bacterium, commande or insect (e.g. European corn borer), preferably schematode or insect (e.g. European corn borer), preferably construct containing the gene into the genome of the plant. The gene is construct containing deme expression in a plant, in response to a stimulus such as infection with a pathogen, damage from a pathogen, construct as infection with a pathogen, damage from a pathogen, constic acid or expression of a gene encoding oxalic acid or expression of a gene encoding oxalic acid oxidase. The constitution in the genes are also useful for stem-preferred regulation of gene expression in a plant. The genes are useful in agriculture, particularly in the chacing of crop plants with improved agronomic traits, for modifying abscisic acid (ABA) metabolism and for modifying manno acid transport and content in plants. The present sequence represents a glutamic acid rich sich such is special form Synechocystis sp. The protein is used in the characterisation of sunflower AAP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           46;
                                                                            Defence-related signalling gene; sunflower; neoxanthin cleavage enzyme; NCE; amino acid permease; AAP; glutamic acid rich protein; GRP; pathogen resistance; abscisic acid metabolism.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Novel isolated defence-related signalling gene isolated from sunflower encoding neoxanthin cleavage enzyme, amino acid permease or glutamic acid-rich protein useful for increasing resistance of plant to a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ::|:: | | |:: :| | |:: :| | |:: :| | |:: :| | |:: :| | |:: :| |:: :| |:: :| |:: :| |:: :| |:: :| |:: :| |:: :|
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        uery Match 5.9%; Score 301.5; DB 22; Length 1749; est Local Similarity 19.0%; Pred. No. 1.1e-07; atches 225; Conservative 117; Mismatches 402; Indels 439; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                17 VSLLAATASFAVAAALGVLPIAITASPAHAATTQPYTWSNVAIGGG--------62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ::||| ::||:
23 lallacsssfsg----nvlaqnitpapdgtgttvdaggngfnigggslsgdggnlfhslg 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------GFVDGIVFNEGAPGILYVRTDIGGMYRWDA 92
Glutamic acid rich protein-like protein amino acid sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---NGRWIPLLDWVGWNNWGYNGVVS--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Example 1; Page 128-135; 135pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           .;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ľ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (PION-) PIONEER HI-BRED INT INC. (CURA-) CURAGEN CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hu X,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     <del>..</del>
_
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18-AUG-1999; 99US-0149656.
23-MAY-2000; 2000US-0206405.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                17-AUG-2000; 2000WO-US22961
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bidney DL, Crasta OR,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2001-211215/21.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          63 -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1749 AA;
                                                                                                                                                                                                                                                      Synechocystis sp.
                                                                                                                                                                                                                                                                                                                                           WO200112801-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               93 A-----
                                                                                                                                                                                                                                                                                                                                                                                                                           22-FEB-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   pathogen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οy
```

AAY13492 standard; Protein; 1426 AA.

AAY13492;

```
The invention relates to a recombinant cellulase active protein free of proteinases of native thermophilic and alkalinophilic origin, comprising the truncated sequences Cell B5, Cell B1, Cell E1/21, Cell 6 or Cell E3/B5, or a stability region from one of the defined full-cel for Cell B1/B5, or M1426, and Cell B4/5 extends from amino CC acid A1011 to P1424 or K1425 or N1426, and Cell B4/5 extends from amino CC acid K635 to N1426 in the sequence shown in AAY1349; Cell E1 extends from amino acid K33 to N1426 in the sequence shown in AAY1349; Cell E1/21 extends from amino acid K1331 to CK K1751 and the stability region extends from amino acid V1233 to CK K1751 and the stability region extends from mino acid E482 to G635 in the sequence shown in AAY13493; Cell E3/B5 is shown in AAX13494. The new CC enzymes are useful in laundry detergent compositions to prevent or remove staining, backstaining or graying, for use on cellulosic materials including oction-containing fabrics. They are especially useful for processing of textiles where cellulose breakdown is required. The new crucated enzymes show reduced redeposition of dye compared to using contrurcated enzymes show reduced redeposition of dye compared to using contrurcated enzymes show reduced redeposition of dye compared to using
                                                                                                                                              Cellulase; proteinase; truncated; Cel B5; Cel B4/5; Cel E1, Cel E1/2; Cel 1/2/3; Cel 6; Cel E3/B5; laundry detergent; stain-removing; cotton-containing fabric; stonewashing.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             835 SRSPSPSASPSSSPSSSPSSSPSSSPSSPVSG-GVKVQYKNNDSAPGDNQIKPGL 893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          894 OVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAIGCGNIRASFGSVNPATPTADT 953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New truncated cellulase proteins, useful in detergents and for producing 'stonewashed' denim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 5.4%; Score 276; DB 20; Length 1426; Best Local Similarity 42.3%; Pred. No. 1.8e-06; Matches 52; Conservative 29; Mismatches 40; Indels 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        , Bergquist PL, Daniels RM, Farrington GK;
Morgan H, Williams DP;
                                                                                                          Truncated cellulases Cel B4/5 and Cel B5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 7; Page 34-37; 65pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                      (CLRN ) CLARIANT FINANCE BVI LTD.
                                                                                                                                                                                                                                                                                                                                              98EP-0810919.
                                                                                                                                                                                                                                                                                                                                                                                  97US-0932571.
                                                                        30-JUL-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1999-315403/27.
N-PSDB; AAX55661.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 1426 AA;
                                                                                                                                                                                                                            Unidentified.
                                                                                                                                                                                                                                                                                                                                              15-SEP-1998;
                                                                                                                                                                                                                                                                                                                                                                                  19-SEP-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ||:
494 yle 496
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Anderson P,
Gibbs MD, M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    954 YLQ 956
                                                                                                                                                                                                                                                                  EP921188-A2.
                                                                                                                                                                                                                                                                                                      .09-JUN-1999
qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
```

.; ;

RESULT AAY13492

AAW15238 standard; protein; 531 AA.

AAW15238

Dockerin; CelD; CelS; CipA; scaffoldin; cellulose binding domain; chromatographic separation; soluble substrate modification; CBD; multi-enzyme delivery system; animal feed; paper production; plant protection; pest control.

Scaffoldin protein from Clostridium thermocellum.

28-JAN-1998 (first entry)

AAW15238;

WO9714789-A2

24-APR-1997

16-OCT-1996; 17-NOV-1995; 17-0CT-1995;

/label= internal_repeat_element_1

154..306

Location/Qualifiers

Region Region Domain

Key

Clostridium thermocellum.

```
The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PRR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful in medical imaging of sites expressing (II). (I) and (II) are useful in retarting disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity amino acid sequences of data and products dependent on DNA and amino acid sequences of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess
                                                                                                                                                                                                                                                                                                                                                               Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 20; SEQ ID No 57609; 103pp; English.
                                                                                                                                                                                                                                                                           Novel human diagnostic protein #27241.
. ABG27250 standard; Protein; 406 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Tang ŸT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         30-MAR-2001; 2001WO-US08631.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         31-MAR-2000; 2000US-0540217.
23-AUG-2000; 2000US-0649167.
                                                                                                                                                                             18-FEB-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2001-639362/73.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Drmanac RT, Liu C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (HYSE-) HYSEQ INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 406 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-PSDB; AAS91437.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200175067-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          biodiversity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11-0CT-2001
                                                                                         ABG27250;
```

```
This protein sequence represents a scaffoldin based on the Ciph protein of Clostridium thermocellum. The scaffoldin protein is used in a novel composition that comprises at least 2 enzymes non-covalently bound to a peptide backbone (i.e. present sequence). The scaffoldin comprises a number of internal repeating units and at least one cellulose binding domain (CBD) The CBD may be altered to modify its affinity for cellulose, which may be desirable where cellulose binding would be disadvantageous. The enzyme is bound to the scaffoldin by a dockerin region of the enzyme, which binds to a repeating element of the scaffolding. The dockerin is preferably CCBs (AAW15237) or Cell (AAW15236). The composition can be used in reducing allergenicity, producing synergistic effects, and facilitating selective modification of substrate. By taking advantage of the cellulose binding domain of the complex, the complex could be immobilised for use in chromatographic separations or for soluble substrate modification. By adding the caffolding domain, it is possible to recover enzymes, or to quantify the amount of an enzyme delivery system which could be used in the food industry, in food processing, animal feed, textiles, bioconversion, pulp and paper production, plant protection and best control, as a wood preservative, topical lotions, and biomass conversions.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Composition with enzymes non-covalently bound to a peptide backbone - used as a multi-enzyme delivery system, e.g. in food processing,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 10; Fig 6; 29pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   textiles and pest control
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1997-245106/22.
```

1;

Gaps

2;

Length 406; Indels 844 PSPSSSPS--PSSSPSSPSPTPSSSPV 869

97

ò g ŏ

Query Match 5.3%; Score 274; DB 22; Best Local Similarity 63.6%; Pred. No. 5.8e-07; Matches 56; Conservative 2; Mismatches 28;

Query Match

Ward M, Xia

Bott RR, Clarkson KA, Fowler T, Liu C,

(GEMV) GENENCOR INT INC.

95US-0559968. 95US-0005701. 96WO-US16485

```
: : | | | :| ||||:|||| :||||| 377 npsdttnsinpqfkvtntgssaidlskltlryyytvdgqkdqtfw-cdhaaiigsngsyn 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAI------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   933 -GCGNIRASFGSVNPATPTADTYLQ 956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           436 gitsnvkgtfvkmssstnnadtyle 460
                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   881
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                             20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Multimer; enzyme; complex; protein-protein interaction; dockerin domain; cohesin domain; catalytic subunit; scaffold subunit; SdbA; synergistic; cellulosome integrating protein; scaffoldin dockerin binding protein.
                                                                                                                                                                                                                                                                                                                                                       279
                                                                                                                                   618
                                                                                                                                                                                    629
                                                                                                                                                                                                                                 ------GVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG 706
                                                                                                                                                                                                                                                        136 gkvtgsvgttveipvyfrgvpsk----giancdfvfrydpnvleiigidpgdiivdpnpt 191
                                                                                                                                                                                                                                                                                KSAPGSSYP----AVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGD--H 760
                                                                                                                                                                                                                                                                                                    ksfdtalypdrklivflfaedsg-tgay--------aitkdqvf 226
                                                                                                                                                                                                                                                                                                                                820
                                                                                                                                                                                                                                                                                                                                                                               SSSPSPSPSPSPSPSRSPSPSASPSPSSSPSPSSSPSPSSSPVSGGVKVQYKNN 880
                                                                                                                                                                                                                                                                                                                                                                                             ::: :|: | : :|: | : :|: | tntatptksatatptr------psyptntptntptntpantpvsgnlkvefyns 323
                                                                                                                                                                                                                                                                                                                                                                                                                              DSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAI----- 932
                                                                                                                                                                                                                                                                                                                                                                                                                                              :: | | | | :| |||||::||| :||||:|||324 npsdttnsinpgfkvtntgssaidlskltlryyytvdggkdgtfw-cdhaaiigsngsyn 382
                                                              Gaps
                                                                                   508 VPSTIFTSPVFTTG---TSVDYAELNPSIIVR----AGSFDPSSQPNDRHVA--FSTDGG
                                                                                                           2 vpskgmancdfvlgydpnvlevtevkpgsiikdpdpsksfdsaiypdrkmivflfaedsg
                                                                                                                                  559 KNWFQGSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQI
                                                                                                                                                          r------gtyaitgdg------vfat-----ivatv
                                                                                                                                                                                   RSDRVNPKTFYALSNGTFYRS------TDGGV-----TFQPVAAGLPSSGAV---
                                                                                                                                                                                                         81 ksaaaapitl--levgafadndlveisttfvaggvnlgssvpttgp---nvpsdgvvvei
                                                                                                                                                                                                                                                                                                                              761 ANLRRVYIGTNGRGIVYGDIGGAPSGSPSPSVSPSASPSLSPSPSSSSPSPSSSSSP
                                                                                                                                                                                                                                                                                                                                                     akiratvkssapgyitfdevggfad-----ndlveqkvsfidggvnvgnatptkgatp
                                                             155;
                                      Length 531;
                                                             Indels
                                      DB 18;
                                   ; Score 261.5; DB 18;
; Pred. No. 3.5e-06;
74; Mismatches 159;

 C. thermocellum cellulosome integrating protein.

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAW43108 standard; Protein; 1853 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             383 gitsnvkgtfvkmssstnnadtyle 407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -GCGNIRASFGSVNPATPTADTYLQ 956
                                     5.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  96FR-0005854
                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Clostridium thermocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Leibovitz E;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (INSP ) INST PASTEUR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1998-011569/02.
                                                Similarity
 Ä
531
                                   Query Match
Best Local Simil
Matches 117; C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FR2748479-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10-MAY-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          10-MAY-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                16-0CT-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 14-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Beguin P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAW43108;
                                                                                                                                                                                                                                                                                                                                                                                                                              881
                                                                                                                                                                                  619
                                                                                                                                                                                                                                                                                707
                                                                                                                                                                                                                                                                                                      192
                                                                                                                                                                                                                                                                                                                                                                              821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAW43108
S
                                                                                                          g
                                                                                                                                                         g
                                                                                                                                                                                ò
                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                   à
                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                              οχ
                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
Multimeric protein, especially enzymatic, complexes are held together by protein-protein interactions between domains designated dockerins and cohesins, which are found on the catalytic and scaffold subunits respectively. An example of such a complex is the cellulose degrading protein complex from clostridium thermocellum, known as the cellulosome. This complex comprises around 15 proteins including endoglucansses, cellobiohydratases, hemicallulases, e.g. xylanases or lichinases, which interact with a central "scaffold" protein designated the cellulosome interact with the cips subunit via conserved 23 amino acid dockerin compared to the subunit via conserved 23 amino acid dockerin compared to the isolation of proteins binding to a novel dockerin type domain found in the C-terminal portion of Cip. The new domain is designated a type II dockerin domain (as compared to the type I domain found on the catalytic subunits of the cellulosome). The type II dockerin domain has some sequence similarity to the type I dockerins but is unable to bind type I cohesin domains. The type II dockerin domain has some sequence similarity to the type I dockerins but is unable to bind type I cohesin domains. The type II dockerin type II dockerin type II dockerin sequence presented here is the cellulosome integrating protein (CIP-A) to which the proteins SdbA, OlpB and ORF2P (AAM43107) can bind at the C-terminal dockerin type II dockerin and cohesin domains can be used in complexes, especially enzyme complexes, to potentiate their used in complexes, especially enzyme complexes, to potentiate their
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   559 KNWFQGSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQI 618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 134 ksaaaaapit1--levgafadndlveisttfvaggvnlgssvpttgp---nvpsdgvvvei 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      189 gkvtgsvgttveipvyfrgvpsk----giancdfvfrydpnvleiigidpgdiivdpnpt 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KSAPGSSYP----AVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGD--H 760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----aitkdgvf 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         280 akiratvkssapgyitfdevggfad-----ndlvegkvsfldggvnvgnatptkgatp 332
                                                                                                    Cellulase proteins with cohesin or dockerin type II domains - useful for potentiating the activity of multiprotein enzyme complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              508 VPSTIFTSPVFTTG---TSVDYAELNPSIIVR----AGSFDPSSQPNDRHVA--FSTDGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --GVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    761 ANLRRVYIGTNGRGIVYGDIGGAPSGSPSPSVSPSASPSLSPSPSSSSPSPSPSPSPSPSSSSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       619 RSDRVNPKTFYALSNGTFYRS-----TDGGV-----TFQPVAAGLPSSGAV--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 19; Length 1853;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         74; Mismatches 159;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5.1%; Score 261.5; DB 333.2%; Pred. No. 1.4e-05
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          245 ksfdtaiypdrkiivflfaedsg-tgay-----
                                                                                                                                                                                                                                                         Disclosure; Page 47; 60pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches 117; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1853 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
N-PSDB; AAT86625
```

20;

1664 AA;

Sequence

g

```
Multimeric protein, especially enzymatic, complexes are held together by protein-protein interactions between domains designated dockerins and cohesins, which are found on the catalytic and scaffold subunits respectively. An example of such a complex is the cellulose degrading protein complex from Clostridium thermocellum, known as the cellulosome. This complex comprises around 15 proteins including endoglucansaes, cellobiohydratases, hemicallulases, e.g. xylanases or lichinases, which interact with a central "scaffold" protein designated the cellulosome interact with the Cip subunit via consarved 23 amino acid dockerin domains. Cip has been shown to contain 9 copies of a cohesin domain. The invention relates to the isolation of proteins binding to a novel domain is designated a type II dockerin domain found in the C-terminal portion of Cip. The new domain has some sequence similarity to the type II dockerin domain has some sequence similarity to the type II dockerin domains. The sequence presented here is an example of a protein which binds the novel type II dockerin domains in the N-terminal portion of which the first domain and is the product of the oldby gene. The novel type II dockerin domains in the N-terminal portion of which the first domain (amino acid residues 28-19) is thought to bind cipa. The novel type II dockerin and cohesin domains can be used in complexes, especially enzyme complexes, to potentiate their catalytic actions in a synergistic manner.
                                                                                                                                                                          Multimer; enzyme; complex; protein-protein interaction; dockerin domain; cohesin domain; catalytic subunit; scaffold subunit; SdbA; synergistic; cellulosome integrating protein; scaffoldin dockerin binding protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Cellulase proteins with cohesin or dockerin type II domains - useful for potentiating the activity of multiprotein enzyme complexes
                                                                                                                                                                                                                                                                                                                                                          /note= "cohesin type II domain"
409..565
/note= "cohesin type II domain"
607..763
/note= "cohesin type II domain"
                                                                                                                                                                                                                                                                                                                          II domain"
                                                                                                                                                                                                                                                                                                28..192
                                 AAW43106 standard; Protein; 1664 AA.
                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Claim 7; Page 31-39; 60pp; French
                                                                                                                                      C. thermocellum OlpB protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               96FR-0005854
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            96FR-0005854
                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                           ..363
                                                                                                                                                                                                                                                 Clostridium thermocellum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Leibovitz E;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (INSP ) INST PASTEUR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1998-011569/02.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-PSDB; AAT86623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FR2748479-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10-MAY-1996;
                                                                                                      16-0CT-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10-MAY-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          14-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Beguin P,
                                                                     AAW43106;
                                                                                                                                                                                                                                                                                                       Domain
                                                                                                                                                                                                                                                                                                                                             Domain
                                                                                                                                                                                                                                                                                                                                                                               Domain
                                                                                                                                                                                                                                                                                                                                                                                                                  Domain
œ
                AAW43106
```

```
44;
                                                                                                          156 QITPLPFKLGGNMPGRGMGERLAVDPNNDNI--LYFGAPSGKGLWRSTDSGATWSQMTNF 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     740
                                                                               100 LLDWVGWNNWGYN----GVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATW 155
                                                                                                                                                                                                      ----ikienmknfagyqlnikydptmleaieletgsaiakrtwpvt--ggtvlqsdny 275
                                                                                                                                                                                                                                                                                                                                                                                                                  324 TSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQI-SWWPDTIIFRSTDGGAT 382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 383 WTRIWDWTSYP-----NRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAIDPF 436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       573 epteepvptetpvdptptvteepvpselpdsyvimeldktk------vkegdviia 622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      tirvnniknlag-----yqigi-------kydpkvleafnietgdp 656
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----NGTFYRSTD---GGVTFQPVAAGLPSSGAVGVMF--HAVPGKEGDLWLAAS 678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                        -----gkttavandv-gagiinfaeaysnltkyret---gvaeetgiigkigfrv 321
                                                                                                                                                                                                                                                                                                                                270 RDGGAT-----WQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSV 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                        371 epgeepteepvp-tetpvd-------ptptvteepvpselpdsyvimeld--kt 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NSDRMLYGT--GATLYATNDLTKWDSGGQIHIAPMVKGL---EETAVNDLISPPSGAPLI 491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SGLYHSTNGGSSWSAITGVSSAVNVGFG-------KSAPGSSYPAVFVVGTIG 724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               725 GVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDIGGAP 784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        741 -idgtymld----wyl-----nrisg--------yvviqpap 764
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 322 lkagstairfedttampgaiegty----mfdwyg-----enikgysvvqpgeivaege
                                                                                                                                                                                                                                                 214 PDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPV----FWS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        514 giigkigfrvlkagstairfedttampgalegtymfdwygenikgysvvqpgeivaege-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----tgiig--nigfrvlkaedttirfeelesmpgs-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 492 SALGDLG------GFTHADVTAVPSTIFTSPVFT-TGTSV-DYAELNPSIIVRAGSF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DPSSQP--------NDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALS ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               657 idegtwpavggtilknrdylptgvainnvskgilnfaayyvyfddyreegksed-----
                                          Indels 292;
  Length 1664;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           872
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |:|| |:|| |:|| |:|| || ||| || || ptpsdeptpsdeptpsdeptpsdeptpsdeptpsdeptpsdeptpsdeptpsdeptpsdeptpsdeptpsdeptpsd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PSPSASPSSSSPSSSSPSSSPSSSPSSPVSGGVKVQYKNNDSAPGD
                                        316;
4.8%; Score 244; DB 19;
20.8%; Pred. No. 0.00011;
iive 96; Mismatches 316;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ABB71513 standard; Protein; 950 AA.
                                      Matches 185; Conservative
                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ABB71513;
  Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  838
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           825
                                                                                                                                                                                                         224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ABB71513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 S X B
                                                                                                                     g
                                                                                                                                                                Q
                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ğ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

ω

```
785
                       ò
                                                 8
                                                                          ò
                                                                                                g
                                                                                                                            ò
                                                                                                                                                   g
                                                                                                                                                                            ò
                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                             29;
                                                                                                                                                                                                                                                                                                                                                                                                                       is
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sequences (ABLO1840-ABL16175) and the encoded proteins (ABB57737-ABB72072).
The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                             New isolated nucleic acid detection reagent for detecting 1000 \, \, \mathrm{or} more genes from Drosophila and for elucidating cell signalling and cell-cell
                                                                                                                                                                                                                                                                                                                                                                                                      The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention useful in developmental biology and in elucidating cell signalling and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    405 AEPWLTFGVQPNPPV-----PSPKLGWMDEAMAIDPF---NSDRMLYGT--- 445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           461 appsssygappgapvssylppasrpskpsssyg----apsvssfvplpsapstnygapsk 516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----GATLYATNDLTKWDSGGQIHIAPMVK--GLEETAVNDLISPPS----GAPLISALGD 496
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                572 sgsf----saapsslysap--skgss-----ggsf--gsapsssysapsas 609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             672 DLWLAASSGLYHSTNGG-----SSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIG 724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                       insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                               interactions in higher eukaryotes for the development of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       tqs1qssgyssgpsssyea----pvappsssygapsssfqp-isppsssygapssgsgs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          557 GGKNWFQGSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAA----SQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ansg---gsypsapsssysapssssss----ggp----yasapssysapssgsnsg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           developmental biology; cell signalling; insecticide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         63; Mismatches 192; Indels 169;
                                                                                                                                                                                                                                                                                                                                                                               Disclosure; SEQ ID NO 41331; 21pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 950;
                                     Drosophila melanogaster polypeptide SEQ ID NO 41331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4.5%; Score 229.5; DB 22; 25.4%; Pred. No. 0.00033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -
-
-
                                                                                                                                                                                                                                                                  Myers EW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ..
=
=
=
                                                                                                                                                                                                                                                                  Li PWD,
                                                                                                                                                                                                   23-MAR-2000; 2000US-191637P.
11-JUL-2000; 2000US-0614150.
                                                                                                                                                                           23-MAR-2001; 2001WO-US09231
           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          144; Conservative
                                                                                                  Drosophila melanogaster.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          <del>-</del>-
                                                                                                                                                                                                                                                                 Venter JC, Adams M,
                                                                                                                                                                                                                                                                                        WPI; 2001-656860/75.
N-PSDB; ABL15616.
                                                                                                                                                                                                                                        (PEKE ) PE CORP NY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            950 AA;
                                                                                                                          WO200171042-A2.
                                                                          pharmaceutical
                                                                                                                                                                                                                                                                                                                                                         interactions
           26-MAR-2002
                                                                                                                                                   27-SEP-2001
                                                              Drosophila;
                                                                                                                                                                                                                                                                                                                                                                                                                                               cell-cell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local
Matches 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       497
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
```

õ g ò g ŏ ð

```
Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic polypeptides. The proteins and their associated DNA sequences are used in the treatment, prevention and diagnosis of medical conditions caused by P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and osteomyelitis), uvethis and endophthalmitis. P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis; useltis; endophthalmitis; bone; joint; central nervous system; ELISA; inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay; dermatological; osteopathic; neuroprotectant.
                                                                                                                                                                                                                                                                                                                                                                                                                                        685 ssysapspg---snsggpypaapsssysaps--psansggpyasapsssysapssssnsg 739
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   polypeptides and nucleic acids useful for diagnosing infections, especially useful for
                                                                                                                                                                                                                                                                                                              768 sssysapssslssggpypsapsssyaapspssnsggpypaapsnsysapiappsssygap
                                                                            725 GVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDIGGAP
                                                                                                                                                                                                                                                                                                                                                                                       835 SRSPSPSASPSSSPSSSPSSSPSSPSPTPSSSPVSGGVKVQYKNNDSAPGDNQIKPGLQ
                                                                                                                                                                                                                                  Bhatia A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Propionibacterium acnes immunogenic protein #5242.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Wang SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Skeiky YAW, Persing DH, Mitcham JL, Wang S
L'maisonneuve J, Zhang Y, Jen S, Carter D;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQ ID No 5541; 1069pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     884 --stgsghsfssapsssysappaggsss 909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             895 VVNTGSSSVDLSTVTVRYWFTRDGGSST 922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Æ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAU44346 standard; Protein; 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21-APR-2000; 2000US-199047P.
02-JUN-2000; 2000US-208841P.
07-JUL-2000; 2000US-216747P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      20-APR-2001; 2001WO-US12865.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     27-FEB-2002 (first entry)
                                                                                                                                                     740 gpyaaapsss-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Propionibacterium acnes vaccinating against and treating acne vulgaris -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Propionibacterium acnes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2001-616774/71.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (CORI-) CORIXA CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-PSDB; AAS59522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            07-JUL-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-2001.
```

22;

AA;

551

```
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              X S
                                                                                                                  Dp
                                                                                                                                                              g
                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                   d
                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                     δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                           음
                                                                                                                                         ò
                                                                                                                                                                                      á
                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                             δ
                                                                                                                                                     Ä
                                              printed
downregulate expression and activity of P. acnes polypeptides and therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        The present sequence is the corrected version of the incorrect Bacillus lautus (NCIMB 40250) endoglucanase Endo 3A described in W09110732. Endo 3 can be used in novel method of forming localised colour density variation on the surface of a dyed cellulosic fabric. The method comprises agitating the fabric in an aqueous medium (PH 6.5 to 9.0) containing a family 5 cellulose, e.g. the present sequence, which can hydrolyse p-nitrophenyl-beta abrading agent or cellulose having abrading activity. Each cellulase displays 30 % or more of its maximum activity at pH 7. The process is useful to provide a stone washed look to blue jeans without back staining.
                                                                                                                                                    5
                                                                                                                                                                                                                                                                                                                                                                                                                      wash;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         dyed
                                                                                                                           Length 455;
                                                                                                                                                                                                                                                                                                                                                                          3A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Localised variation of colour density in the surface of a dyec cellulosic fabric – uses cellulase compsn. able to hydrolyse
                                                                                                                                                                                                                                                                                                                                                                      Corrected Bacillus lautus (NCIMB 40250) endoglucanase Endo
                                                                                                                                                                                                                                                                                                                                                                                            Endoglucanase; Endo 3A; formation; localised; variation; colour density; surface; dye; fabric; family 5; cellulose; hydrolysation; p-nitrophenyl-beta-1,4-cellobioside; stone vblue jeans; back stalning.
                                                                                                                                                    Indels
                                                                                                                                                                                      DB 22;
                                                                                                                                                    12;
                                                                                                                           4.4%; Score 227.5; DB 2
34.8%; Pred. No. 0.00018;
tive 41; Mismatches 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Toft AH;
                                                                                                                                                                                                                                   Disclosure; Pages 15-17; 23pp; English.
                                                                                                                                                                                                                     847 SSSPSSSPSSSPSPTP----SSSPVS 870
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            p-nitrophenyl -beta-1,4-cellobioside
                                                                                                                                                                                                                                                                                                      A.
                                                                                                                                                                                                                                                                                                    AAW18790 standard; protein; 551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Schulein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       96WO-DK00364.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              95DK-0000993
                                                                                                                           Query Match 4.4%
Best Local Similarity 34.8%
Matches 31; Conservative
                                                                                                                                                                                                                                                                                                                                               18-NOV-1997 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (NOVO ) NOVO-NORDISK AS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1997-192888/17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Onishi M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Bacillus lautus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO9709410-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       03-SEP-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08-SEP-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13-MAR-1997.
                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                           AAW18790;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Fich M,
                                                                                                                                                                                                                                                                                        AAW18790
                                                                                                                                                                                                                                                                                                                            g
 ò
                                                                                                                                                                                               ద
                                                                                                                                                                                                                      à
```

```
310
                                                                                                                                                                                                                                                                                    ------pnviyeian----epngnvswadvksy 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   750 GNW-------GOAITGDHANLRRVYIGTNGRGIVYGDIGGAPSGSPSPSVSPS 795
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               845
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | :|| : :: | : : | :| 311 aptnls-atagnaqvsltwnavsgatsytvkrattsggpytnvatgvtatsytntgltng 369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          54 ginvfraamytaedgyitdpsvknkvkeavqasidlglyviidwhilsdgnpntykaqsk 113
                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      898 TGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAIGCGNIRASFGSVNPATPTADTYLQ 956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Streptococcus pneumoniae; vaccine; screening; protein antigen; antibacterial; antiinflammatory; meningitis; infection; diagnosis;
                                                                                                                                                                                                                             599 ----AVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF----
                                                                                                                                                                                                                                                                                                                                                                                                                                                    695 TGVSSAVNVGFGKSAPGSSYPAVFVV--GT--IGGVTGAYRSDDCGTTWV-LINDDQHQY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         210 t---yamnkg-----aaifvtewgtsdasgnggpyfpg--skewidflnarkisw
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       255 vnwsladkvetsaalmpgasptggwtdaqlsesgkwyrdqirqatggg-sgnpt---apa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          846 -----PSSSPSSSPSSSPSSPTPSSSPVS-GGVKVQYKNNDSAPGDNQIKPGLQVVN
                                                                                                                                                                                                                                                                                                                                           647 -QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHS------TNGGSSWSAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               558 GKNWFQG----SEPGGVTTGG-----TVAASADGSRFV---WAPGDPGQPVVY----
                                                                                                                                                                                                                                                                                                                                                                      New Streptococcal protein, useful as a vaccine, for diagnosis of
4.4%; Score 225.5; DB 18; Length 551;
23.8%; Pred. No. 0.00029;
iive 72; Mismatches 166; Indels 127;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Streptococcus pneumoniae type 4 protein sequence #109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ā
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY81609 standard; Protein; 1237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (MICR-) MICROBIAL TECHNICS LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98GB-0016337.
99US-0125164.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99WO-GB02451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                             Best Local Similarity 23.89
Matches 114; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Streptococcus pneumoniae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hansbro PM
                                                                                                                                                                                                                                                                                    114 affgematlygnt-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            pneumococcal disease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2000-195300/17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO200006737-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gilbert CFG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       27-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              27-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24-MAY-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LO-FEB-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY81609;
     Query Match
```

Thu Jul

24; AAY81501 to AAY81679 represent specifically claimed protein sequences isolated from Streptococcus pneumoniae. AAA05407 to AAA05590 represent specifically claimed nucleotide sequences isolated from S. pneumoniae. The sequences have antibacterial and antihilammatory properties. The protein sequences, and fragments of them, are useful as immunogens and/or antigens. The nucleotide sequences can be useful as immunogens and/or antigens. The proteins and nucleotides can be useful for the detection and diagnostis of S. pneumoniae. The protein sequences are also useful for screening an agent capable of antagonising, inhibiting or interfering with the function or expression of the proteins in which the agent is useful for treatment or prophylaxis of S. pneumoniae infection and meningitis. AAA05591 to AAA05614 represent primers used in the exemplification of the present invention. pneumococcal diseases and for screening agents capable of antagonizing or inhibiting expression of the protein 305 350 595 386 655 442 530 sastsastsisasstvvgsqtaaateatakkveedrkkpasdyvasvtnvnlgsyakrrk 167 487 775 Gaps tstsasastsasasastsasasasisasesaststsasesastsasesastsas asastsasasastsasastsasastsasasastsasasastsasasasts STDSGATWSOMTNFPDVGTYIANPTDTTG------YOSDIOGVVWVAFDK----SSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYI 168 rsvdsiegllasi-----knaavf----sgnt--ivngapa-----inaslni 303 ATSNTGGPYDGSSGD-----VWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQH 205 akset-kvytgegvdsvyrvpiyyklkvtn------dgskltfty---tvtyvn PNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQP pktndlgnissmrpgysiynsgtstqtmltlgsdlgkpsgvknyitdkngrqvlsyn---416 NPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEE 476 TAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIV -----gltss-----wtvpi--tgtdtsf------536 RAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFVWAPGDPGQP 351 ---tftpyaartdri-----ginyfng-----gkvvessttsgsl-----sgs 596 VVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVAAGLPS 387 kslsvsasgsasasastsasasasts---asasastsasasastsasvsastsasasas SGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYP tsasa------sastsasesastsasasasasastsasasastsasastsase 716 AVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGI sastsasastsasesastsas----asastsasas---astsasass --- SSPVSGGVKVQYKNNDS 214; Length 1237; Score 219; DB 21; Length 12 Pred. No. 0.0016; 6; Mismatches 323; Indels RSPSPSASPSPSSSPSSSSPSSPSPTPS--English. 86; 4.3%; 21.2%; Claim 1; Page 83; 108pp; Conservative Similarity 1237 AA; 168; Query Match Best Local S Matches 168 Sequence 200 108 244 356 249 306 334 959 443 488 836 591 g ò a δ g ò g δ 엄 ογ g ò 8 Dp 셤 g a ò ç ò ð ò

```
ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ဌ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       the encoded proteins (AA000010-PA013910) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The production of other optoxides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities of stem cell growth factor activity, haematopolesis regulating activity, tissue growth factor activity, immunomodulatory activity and activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                            gene therapy;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           The invention relates to human polynucleotides (AAI79941-AAI93841) and
APGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAIGCGNIRASFG 942
                              651 asasastsasvsastsasasastsasastsasesastsasasastsasasastsasasas 710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          vaccine; peptide therapy; stem cell growth factor; haematopolesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorders; arthritis; inflammation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Isolated nucleic acids and polypeptides, useful for preventing diagnosing and treating e.g. leukaemia, inflammation and immune
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 20; SEQ ID NO 16016; 1399pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                            cell differentiation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4.2%; Score 218; DB 22;
54.7%; Pred. No. 9.3e-05;
tve 2; Mismatches 33;
                                                                                                                                                                                                                                                                                                            proliferation;
                                                                                                                                                                               AA.
                                                                                                                                                                                                                                                                            Human polypeptide SEQ ID NO 16016.
                                                                                                                                                                            AAO02124 standard; Protein; 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Liu C, Drmanac RT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2000US-0515126
2000US-0577409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 54.7%;
Matches 47; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              26-FEB-2001; 2001WO-US04927
                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                              cell
                                                                                            711 astsasesast 721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2001-514838/56.
N-PSDB; AAI82055.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (HYSE-) HYSEQ INC
                                                               SVNPATPTADT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           88 AA;
                                                                                                                                                                                                                                                                                                               cytokine;
                                                                                                                                                                                                                                                                                                                                                                                                                             WO200164835-A2.
                                                                                                                                                                                                                                                                                                                                                                                                Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-FEB-2000;
18-MAY-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           inflammation
                                                                                                                                                                                                                                             06-NOV-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                             07-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             disorders
                                                                                                                                                                                                            AA002124;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           rang YT,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                               13
883
                                                                                                                                               RESULT
                                                                                                                                                                                                              q
                                                               ŏ
                                                                                            g
```

= = =

_ _ _ _ _

= = -

ð

8 ò 셤 AAB74609

```
(ABB57737-ABB72072).
The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL36511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   696 sgssg-----lgnnvnvkgdt 729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        84 I---GGMYRWDAANGRWIPLLDWVGWNNWGYNGVV--SIAADPINTNKVWAAVGMYTNSW 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     --LWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSSLGQASKT 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 255 IFVGVADP --- NNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPY 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                312 DGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDT 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         24 ASFAVAAALGVLPIAITASPAHAATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTD 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 494 anvpvgaslpvls-----td
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              575 ssssgs---ssgvgnnfnlapsvqeig------dsntvndlinanvpvgasl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             524 vlvngtgntfnrlvgaripvl-----anapilssinsgssgtsgssssgssssgss
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      139 DPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDN-ILYFGAPSGKG-
                                                         Drosophila; developmental biology; cell signalling; insecticide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ID NO 39117; 21pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 4.1%; Score 213; DB 22; Length 1 Best Local Similarity 20.0%; Pred. No. 0.0036; Matches 204; Conservative 135; Mismatches 337; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          pvlssitgsgftgntlnggsdv---lingtgnt---
                                                                                                                                                                                                                                                                                                                                                                             Myers
                    melanogaster polypeptide
                                                                                                                                                                                                                                                                                                                                                                             Li PWD,
                                                                                                                                                                                                                                                                            2000US-191637P.
2000US-0614150.
                                                                                                                                                                                                                                       23-MAR-2001; 2001WO-US09231
                                                                                                                 Drosophila melanogaster.
                                                                                                                                                                                                                                                                                                                                                                             Venter JC, Adams M,
                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2001-656860/75.
N-PSDB; ABL14878.
                                                                                                                                                                                                                                                                                                                                      (PEKE ) PE CORP NY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Disclosure; SEQ
                                                                                                                                                        WO200171042-A2.
                                                                                  pharmaceutical
                                                                                                                                                                                                                                                                          23-MAR-2000;
11-JUL-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   interactions
                                                                                                                                                                                                27-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     197
    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The present invention describes a human proteophosphoglycan protein designated hPPG-1. Also describes are methods for the preparation and detection of hPPG-1. The present sequence represents the Leishmania major PPG protein which is used in comparison with the hPPG-1 protein, in an example from the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    783 APSGSPS-PSVSPSASPSLS---PSPSSSSPSSPSPS-PSSSSPSSPSPSPSR 836
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         15 apsssssapsasssapsasssapsasssapsasssapsasssapsassss 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  837 SPSPSAS--PSPSSSPSPSSSPSSPSPTPSSSPVSGGVKVQYKNNDSAPGDNQIKPGLQ 894
  61
.;
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 237;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Human phosphoglycan protein and its coded sequence
                                                                                                                                                                                                                                                                                                                                    Human; hPPG-1; PPG; proteophosphoglycan; detection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 43;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (NANF-) NANFANG RES CENT STATE HUMAN GENE GROUP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4.2%; Score 216; DB 22, 48.4%; Pred. No. 0.00036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              15; Mismatches
                                                                                                                                                                                                                                                                                             Leishmania major PPG protein sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ABB70775 standard; Protein; 1357 AA
                                     843 SPSPSSSPSSSSPSSSPSPTPSSSP 868
                                                                                                                                                                             ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Example 2; Page 20; 22pp; Chinese
                                                             ddssssssdddsdddsdddssdddd-
                                                                                                                                                                         AAB74609 standard; Protein; 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11-APR-2000; 2000CN-0115368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11-APR-2000; 2000CN-0115368
                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity 48.48
hes 62; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          26-MAR-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Xiao H, Kang B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2001-183595/19.
N-PSDB; AAF74867.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             237 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        135 sssapsss 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   895 VVNTGSSS 902
                                                                                                                                                                                                                                                                                                                                                                           Leishmania major.
                                                                                                                                                                                                                                                                                                                                                                                                              CN1272542-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                     08-NOV-2000
                                                                                                                                                                                                                                                       22-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                AAB74609;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Li N,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15
                                                                          62
```

RESULT 1

õ g õ 셤 õ g BXXXE

46;

Gaps

Length 1357;

13

豎

LRYVLDISAEPWLTFGVQPNPPVPSPKLG 425	QIHIAPMVKGLEETAVNDLISPP 485 . : : ssqvtstgssnsts 806	TSPVFTTGTSVDYA-ELNPSI 533 : : :: nvpvlskitgtgfannlnqksnv 865	GSEPGGVTTGGTVAASADGSRFV 586 	NAQIRSDRVNPKTFYALSNG 634 ::	GVMFHAVPGKEGDLWLAASS 679 : : pvlggvlntvepvvgdlvggvls 1041	gnsassgantlitgnnnivrlv 1097	LINDDQH747 : : vigdnnrivklinvdlwvlafld 1157	vyg 778 1 vdvdlsvllvvdllshlitnvan 1213	SPSPSPSSSPSSSPSPSPS 831 : : : : : tptygpptnepptnqpptngpt 1273	-PSSSPVSGGVKVQYKNNDSAPG 885 :: : : pptnepptrrppsps 1325		
IIFRSTDGGATWTRIWDWTSYPNRS 	WMDEAMAIDPENSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEETAVNDLISPP ::: :: :: :: inqgsstgsstgsstgsstgsstgsstgssgssgvtstgssnsts			WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNFKTFYALSNG :		GLXHSTNGGSSWSAITGVSSAVNVGFGKSAPGSS	YPAVFVVGTIGGYTGAX-RSDDCGTTWVLINDDQH	OYGNWGQAITGDHANLRRVYIGTNGRGIVYG	DIGGAPSGSPSPSV\$PSASPSLSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPS	PSPSRSPSPSASPSPSSSPSPSSSPSPTPSSSPVSGCVKVQYKNNDSAPG	DA —	pn 1327
372 730	426	486	534	587 925	635 984	680	714	748	779	832	886	1326
Qy Dp	Qy	Qy Db	Qy	Qy Dp	Qy Db	Qy Dp	Oy Dp	Qy Dp	Oy Op	Oy Dp	oy .	අ

Search completed: July 2, 2002, 09:15:39 Job time: 227 sec

3, Appli 2, Appli 4, Appli 5, Appli 5219987

Appli Appli

Appli Appli Appli Appli Appli Appl

```
Anderson, Paige
Glbbs, Moreland
Bergquist, Peter
Daniels, Roy
Morgan, Hugh W.
Williams, Diane P.
INVENTION: Compositions and Methods for
Treating Cellulose Containing Fabrics Using Truncated
Cellulase Enzyme Compositions
                                                                                                                                                                                                      Sequence 4, Appli
              Sequence 6, Al
Sequence 4, Al
Sequence 7, Al
Sequence 2, Al
Sequence 92,
Sequence 93,
Sequence 2, Al
Sequence 2, Al
Sequence 2, Al
Sequence 2, Al
                                                                                                                                                     Sequence 5,
Patent No. 5,
                                                                                                                                                                             Sequence 2,
Sequence 1,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Box 457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STREET: Spring House Corporate Center, P.O. CITY: Spring House
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: 1997US001/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
US-08-217-327-4
US-07-667-276A-6
US-08-395-602A-4
US-08-021-625D-4
US-07-862-58BB-7
US-09-110-517-2
US-08-81B-111-93
US-09-103-429A-3
US-09-103-429A-3
US-08-983-045-2
US-08-983-045-2
US-08-410-784A-5
52198R-7
PCT-US93-08435-2
PCT-US93-08435-2
US-08-642-255-62
US-08-642-255-62
US-08-642-255-62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,574A
FILING DATE: 19-Aug-1998
CLASSIFICATION: <UNKNOWN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/932,571
FILING DATE: September 19, 1997
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                        ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 44:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: Howson and Howson
                                                                                                                                                                                                                                                                                                                    Sequence 44, Application US/09136574A
Patent No. 6294366
GENERAL INFORMATION:
APPLICANT: Farrington, Graham K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER: IBM Compatible OPERATING SYSTEM: DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH: 1751 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELEFAX: 215-540-5818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO: 44:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NAME: Bak, Mary E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELEX: <Unknown>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NUMBER OF SEQUENCES:
  214
2554
1126
1126
1531
1581
303
303
303
303
702
772
772
1160
1163
1164
11064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY: USA
ZIP: 19477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STATE: PA
                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE OF
                                                                                                                                                                                                                                                                                                       US-09-136-574A-44
  2, 2002, 09:11:52; Search time 39.87 Seconds (without alignments) 586.288 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 1, Appli
Sequence 3, Appli
Sequence 10, Appl
Sequence 12, Appl
Sequence 2, Appli
Sequence 2, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 1, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 4, Appli
Patent No. 5202236
Patent No. 5202236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      , Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 9, Appli
Sequence 3, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 44, Appl
Sequence 43, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 40,
Sequence 40,
Sequence 40,
Sequence 40,
Sequence 40,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 47,
Sequence 18,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 2, P
Sequence 2, P
Sequence 1, P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 3,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
                                                                                                                                                                 1 MDRSENIRLTMRSRRLVSLL.....RASFGSVNPATPTADTYLQX
                                                                                                                                                                                                                                                                                                                                                                               /cgn2_6/ptodata/2/jaa/5A_COMB.pep:*
/cgn2_6/ptodata/2/jaa/5B_COMB.pep:*
/cgn2_6/ptodata/2/jaa/6A_COMB.pep:*
/cgn2_6/ptodata/2/jaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/jaa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/2/jaa/PCTUS_COMB.pep:*
                             Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-276-213-3
US-09-198-956-10
US-09-198-955-12
US-07-109-811-2
US-09-109-811-2
US-09-118-574A-47
US-08-213-2888-18
US-08-293-728-2
US-08-293-728-2
US-08-293-58-4
US-08-932-9298-1
US-08-932-9298-1
US-08-939-411-4
US-08-399-411-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-599-654-40
US-08-944-868A-40
US-08-944-423A-40
US-08-944-496-40
                                                                                                                                                                                                                                                       hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PCT-US95-13813-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-760-797A-3
US-08-932-929B-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-616-844-40
                                                                                                                                                                                                                              231628 segs, 24425594 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SUMMARIES
               GenCore version
Copyright (c) 1993 - 2000
                                                                                                                                                                                                                                                                                                                  Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                OM protein - protein search, using sw model
                                                                                                                                                                                                     Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                     Issued_Patents_AA:*
                                                                                                                                                                                                                                                                               Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                    US-09-917-376-1
5135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          424
424
1719
1719
1719
334
334
424
424
                                                                                                                                                                                          BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query
Match 1
                                                                                        July
                                                                                                                                                                                                                                                       oţ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2276
276
2205.5
2205.5
2208.5
208.5
208.5
208.5
199.5
199.5
199.5
185.5
185.5
185.5
185.5
185.5
185.5
185.5
185.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                180.5
180.5
179.5
179.5
179.5
179.5
                                                                                                                                                      Perfect score:
                                                                                                                                                                                          Scoring table:
                                                                                                                                                                                                                                                      Total number
                                                                                                                                                                    Sequence:
                                                                                                                                                                                                                              Searched:
                                                                                                                                                                                                                                                                                                                                                                      Database
                                                                                         Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ٠
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Result
```

~

ò 염 ò

ò

ò

ò

ò

ŏ

ò

ŏ g Óγ g g

ò qq ŏ

ð g ò

```
Fabrics Using Truncated
    938
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     835 SRSPSPSASPSPSSSPSPSSSPSSPTPSSSPVSG-GVKVQYKNNDSAPGDNQIKPGL 893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAIGCGNIRASFGSVNPATPTADT 953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Center, P.O. Box 457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 1426,
  888 KVWYANGNLSSPINVLNPKIKIENVGTTAVDLSRVKVRYWYTIDGEATQSV-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                         Treating Cellulose Containing Cellulase Enzyme Compositions
                                                                                                                                                                                                                                                                                                                                 Daniels, Roy
Morgan, Hugh W.
Williams, Diane P.
TITLE OF INVENTION: Compositions and Methods for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         40;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: 1997US001/CIP
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 5.4%; Score 276; DB 4; Best Local Similarity 42.3%; Pred. No. 5.2e-10; Matches 52; Conservative 29; Mismatches 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OPERATING SYSTEM: DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,574A
FILING DATE: 19-Aug-1998
CLASSIFICATION: <UNKNOWN>
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: US 08/932,571
FILING DATE: September 19, 1997
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE DESCRIPTION: SEQ ID NO: 43: US-09-136-574A-43
                                                                                                                                                                                       Sequence 43, Application US/09136574A Patent No. 6294366 GENERAL INFORMATION:
                                                                                                                                                                                                                                                     APPLICANT: Farrington, Graham
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER: IBM Compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH: 1426 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEPHONE: 215-540-9200
                                                                                                                                                                                                                                                                        Anderson, Paige
Gibbs, Moreland
Bergquist, Peter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOLECULE TYPE: No. 6294366e
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEFAX: 215-540-5818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO: 43: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCES: 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Spring House
                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQUENCES: 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM MEDIUM TYPE: Diske
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: Bak, Mary E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           <Unknown>
                                                                      : |:|||
----SVASSINPA 947
                                         934 CGNIRASFGSVNPA 947
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ZIP: 19477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STATE:
                                                                                                                                                                    US-09-136-574A-43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         954 YLQ
                                                                                   939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        894
                                         δ
                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                      56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    246
                                                                                                                                                                                     SVVAET-----AASLAVASVVIKERNSOKAASYLQHAKDLFEFADTIRSDAGYTAATGFY 240
                                                                                                                                                                                                                                                                                                                  LLDWVGWNNWGYNGVVSIAADPINTNKVWAAVGM-YTNSWDPNDGAILRSSDQG----- 152
                                                                                                                                                                                                                                                                                                                                                                                              153 ATWQ-------ITPLPFKLGGNMPGR----GMGERLAVDPNNDILYFGAPSG 194
                                                                                                                                                                                                                                                                                                                                                                                                                                          400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----AHSSWANSMKIPEYHRHILYGALVGGPGSDDSYNDDITDYVQNEVACDYNAG 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----MYQLYGGEPID--DFKAIETPINDEIFVESKF 489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVL-DISAEPWLTFGV 413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   414 QPNPPVPSPKLGW----MDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPM 469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      519
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           629 WGLEPGAATSTPAPTSTPTPTPTPTVTA------TPTPTPTPTPTPTPTGSP--G 674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAAS 579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            708
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           709 VDLSRVKIRYWYTVDGDKPQSAVC-----DW------AQIGASNVT-FNFVKLSSG 752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     635 TFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKE-GDLWLAASSGLYHSTNGGSSWSA 693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 753 V-----SGADYY-LEVGF-SSGAQQLQ----PGKDTGDIQVRFNKNDWSNYNQADDWSW 800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           754 QAITGDHANLRRVYIGTNGRGIVYGDIGGAPSGSPSVSPSASPSLSPSPSSSSPSPS 813
                                                                                                                                                                                                                                                                      241 TSGGFIDDLGW---AAVWLYIATNDSSYLTKAEELMSEYANGTNTWTQCWDDVRYGTLIM 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    580 ADGSRF---VW--APGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNG 634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWG 753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      824
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         841
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAIG 933
                                                                                                      Gaps
                                                                                                                                                ----TTQP---YTWSNVAI 59
                                                                                                                                                                                                                                 ---IGGMYRW-----DAANGRWIP 99
                                                                                                                                                                                                                                                                                                                                              LAKITGKEL--YKGAVERNLDHWTDRITYTPKGMAYLTGW----GSLRYATTAAFLACVY
                                                                                                                                                                                                                                                                                                                                                                                                                                      352 ADWSGCDSNKKTKYLNFAKSQIDYALGST -- GRSFVVGFGINYPQHPHHRN-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  195 KGLWRSTDSGATWSQMTNFPD-----VGTYIANPIDTTGYQSDIQGVVW--VAFDKSSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIAT --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTAN-DYF------GYSGLTIDRQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  490 GNSQGP---NYTEVISYIYNRTGW----PPRVTDKLSFKYFIDLTELIQAGYS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -PDVVKVDT-----YYIEGGKISGPYVWD----KNRNIYYVLVDFSGTK----I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             576 YPGGEVEHKKQAQFKISVPQGYPWDPTN-DPSYKGLTSQLEKNKYIAAYDNNN----L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VKGLEETAVNDLISP-----PSGAPLISALGDLGGFTHADVTAVPSTIFT---SPVFT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TGSGVKVLYKNNETSASTGSIRP-------WFK-----IVNGG--SSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ------TPAPISTATPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6.5%; Score 334; DB 4; Length 1751;
llarity 21.2%; Pred. No. 1e-13;
Conservative 125; Mismatches 314; Indels 376;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SLLAATASFAVAAALGVLPIAITASPAHAA--
                                                                                                                                                                                                                               GGGGFVDGIVFNEGAPGILYVRTD-----
                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        453 IVGALAK----
US-09-136-574A-44
                                                           Query Match
Best Local Simi
Matches 219;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   305
                                                                                                                                            18
                                                                                                                                                                                     186
                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                             100
                                                                                                                                                                                                                                                                                                                                                       298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            675
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          694
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       825
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          842
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 874
                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QQ
```

ä

g

```
| :|| : : : | : : | 311 APINLS-ATAGNAQVSLTWNAVSGATSYTVKRATTSGGPYTNVATGVTATSYTNTGLTNG 369
                                                                                                                                                                                                                             370 TTYYVVSASNSAGSSANSAQASATPASGGASTGNLVVQYKVGDTSATDNQMKPSFNIKN 429
                         -----GQAITGDHANLRRVYIGTNGRGIVYGDIGGAPSGSPSPSVSPS 795
                                                                255 VNWSLADKVETSAALMPGASPTGGWTDAQLSESGKWVRDQIRQATGGG-SGNPT---APA 310
                                                                                                                                                                                                   846 -----PSSSPSPSSSPSSPSPSTPSSSPVS-GGVKVQYKNNDSAPGDNQIKPGLQVVN 897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 428 ----DEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPM---VKGLEETAVND 480
                                                                                                                                                                                                                                                                                           898 TGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAIGCGNIRASFGSVNPATPTADTYLQ 956
                                                                                                                                                                                                                                                                                                                2 GGGYW------HTSGREILDAN-----NVPVRIAGINWFGFETCNYVVHG 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4.1%; Score 210.2;
22.2%; Pred. No. 2.9e-06;
ifive 74; Mismatches 198; Indels 207;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          379 GGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWM-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: Thomas, Steven
APPLICANT: Laymon, Robert
APPLICANT: Himmel, Michael
TITLE OF INVENTION: GENE ENCODING FOR THE E1 ENDOGLUCANASE
NUMBER OF SEQUENCES: 6
                                                                                                            CORRESPONDENCE ADDRESS:
ADDRESSEE: National Renewable Energy Laboratory
STREET: 1617 Cole Boulevard .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OPERATING SYSTEM: PC_DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATTORNEY/AGENT INFORMATION:
NAME: 0'Connor, Edna
REGISTRATION NUMBER: 29,252
REFERENCE/DOCKET NUMBER: NREL IR# 94-08
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303)231-1000
                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 3, Application US/08276213
Patent No. 5536655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY: USA
ZIP: 80401-3393
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         : 521 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-terminal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (303)231-1098
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches 137; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEFAX: (303)231-1098
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY: linear
MOLECULE TYPE: pr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE: amino a STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             S
Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Golden
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; FRAGMENT TYPE:
US-08-276-213-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                         750 GNW----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ANTI-SENSE:
                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-276-213-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STATE:
                                                                                                                                                     g
                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                            ò
                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                            No. 59580830 No. 5958083disk of No. 5958083th America, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             114 AFFQEMATLYGNT-------PNVIYEIAN-----EPNGNVSWADVKSY 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   150 AEEVITAIRAIDPDGVVIVGSPTWSQDIHLAADNPVSHSNVMYALHFYSGTHGQFLRDRI 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   695 TGVSSAVNVGFGKSAPGSSYPAVFVV--GT--IGGVTGAYRSDDCGTTWV-LINDDQHQY 749
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHS------TNGGSSWSAI 694
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             210 T---YAMNKG------AAJFVTEWGTSDASGNGGPYFPQ--SKEWIDFLNARKISW 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               558 GKNWFQG----SEPGGVTTGG-----TVAASADGSRFV---WAPGDPGQPVVY---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----AVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 551;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 2;
                                                                                                                                                                                             APPLICANT: Onishi, Masahiro
APPLICANT: Fich, Merete
APPLICANT: Toft, Annette Hanne
APPLICANT: Sh lein, Martin
TITLE OF INVENTION: In Stone Washing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2e-07
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OPERATING SYSTEM: DOS
SOFTWARE: FastSEO for Windows Version 2.0
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. 3.2e
; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4.4%; Score 225.5;
23.8%; Pred. No. 3.2k
tive 72; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE/DOCKET NUMBER: 4492.204-US 
TELECOMMUNICATION INFORMATION: 
TELEPHONE: 212-867-0123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: US/09/033,537A FILLING DATE: 02-MAR-1998 CLASSIFICATION: 008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRIOR APPLICATION DATA:
APPLICATION NUMBER: 0993/95
FILING DATE: 08-SEP-1995
APPLICATION NUMBER: PCT/DK96/00364
FILING DATE: 03-SEP-1996
ATTORNEX/AGENT INFORMATION:
                                                                                                                            Sequence 1, Application US/09033537A Patent No. 5958083
                                                                                                                                                                                                                                                                                                                                                                                            405 Lexington Avenue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME: Green, Reza
REGISTRATION NUMBER: 38,475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IBM Compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    : 551 amino acids
amino acid
EDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS:
LENGTH: 551 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              212-878-9655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:
MEDIUM TYPE: Diskett
                                                                                                                                                                                                                                                                                                                                                       CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                     NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER: IBM CON OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
Matches 114; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   U.S.A.
                                                                                                                                                                            GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                   STREET: 405 Les
CITY: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRANDEDNESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10174
||:
494 YLE 496
                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-033-537A-1
                                                                                                          -09-033-537A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      599
```

31;

ò qq ò g ò

```
17 -----IEGPSKNIWVDHNE-LY 142
                                                                                                                                    683 HSTNGGSSW------SAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYR 731
                                                         523 VNPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLY
                                                                                                                                                                           143 HSLNVDKDYYDGLFDVKRDAEYITFSWNYVHDGWKSMLMGSS----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      408 KGTFVKMSSSTNNADTYLE 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1998-11-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                938 RASFGSVNPATPTADTYLO 956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM: Clostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRIOR FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-198-955A-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQ ID NO 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGTH:
                                                                                                                                                                                                                                                                                                                                    244
                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                           à
                                                                                             g
                                                                                                                                    ò
                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                    δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19;
                                                                                           KI------VAYAGQIGLRIILDRHRPDCSGQSALWYTSSVSEATWISDLQALAQRYKG 150
                                                                                                                                  530 NPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGG------VTTGGTVAA 578
                                                                                                                                                                                                                                      205 SYNGDSYWWG------GNLQGAGQYPVVLNVPNRLVYSAHDXATSVYPQTWF 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                  737 TTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDIGGAPSGSPSVSPSA 796
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    151 NPTVV----GFDLHNEPHDPACWGCGDPSIDWRLAAERAGNAVLSVNPNLLIFVEG--VQ 204
                                                                                                                                                                                                                SADGSRFVWAPGDPGQPVVYAYGFGNSWAASQ-----GVP----ANAQIRSDRVNPKTFY 629
                                                                                                                                                                                                                                                                                         630 ALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGS 689
                                                                                                                                                                                                                                                                                                                                                                       690 SWSAITG--VSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSD------DCG 736
                                                                                                                                                                                                                                                                                                                                                                                                           278 VWLGEFGTTLQSTTDQTWLKTLVQYLRP-----TAQYGADSFQWTFWSWNPDSG 326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       327 DTGGILKDD-----WQTVDTVKDGYLAPI-----KSSIFDPVGA------SA 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                251 --SDPTF-----GYLFNQNIAP 277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            419 ASGARCTAS------YQVNSDWGNGFT-----VTVAVTNSGSVATKTWTVSWTF- 461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     568 GGVTTGGTVAASAD----GSRFVWAPGDPGQPVVYAVGFG-NSWAASQGVPANAQIRSDR 622
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       857 SSSPSPTPSSSPVSGGVKVQYK-NNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFT
                                                       LISPPSGAPLISALGDLG-----GFTHADVTAVPSTIFTSPVFTTGTSVDYAEL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels 125;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 493;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 10, Application US/09198956

Patent No. 6165769

GENERAL INOFMARION:
APPLICANT: Andersen. Lene N.
APPLICANT: Andersen. Lene N.
APPLICANT: Bajornvad, Martin.
APPLICANT: Lange, Niels Erik K.
APPLICANT: Lange, Niels Erik K.
APPLICANT: Bajornvad, Mads E.
APPLICANT: Bajornvad, Mads E.
APPLICANT: Schnorr, Kirk
TITLE OF INVENTION: Licheniformis
FILE REFERENCE: 5377.200-US
CURRENT APPLICATION NUMBER: US/09/198,956
CURRENT APPLICATION NUMBER: 1344/97
EARLIER APPLICATION NUMBER: 1344/97
EARLIER FILING DATE: 1997-11-24
EARLIER FILING DATE: 1997-11-24
EARLIER FILING DATE: 1997-11-24
SEARLIER FILING DATE: 1997-11-24
SEARLIER FILING DATE: 1997-11-24
SEARLIER FILING DATE: 1997-11-24
SOUTHWARE: FastSEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6e-06;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4.1%; Score 208.5;
23.5%; Pred. No. 3.6e
tive 62; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: PRT; ORGANISM: Bacillus licheniformis
US-09-198-956-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --GGNQTITNS--WNA 473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 23.55
Matches 103; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   916 RDGGSSTLVYNCDWAA 931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-198-956-10
                                                       181
                                                                                           66
                                                                                                                                                                                                              579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          462
                                                                                           g
                                                         ð
                                                                                                                                ò
                                                                                                                                                                   g
                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                          D
                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                            Db
                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                     පු
                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
186 SDNYNRIITFHHNWFENLNSRVPSFRFGEGHIYNNYFNKIIDSGINSRMGARIRIENN-- 243
                                                                                                                                                                                                                           295 NVDNVKSIVKQ---NAGVGKIQRRPPTPTPTSPPS---ANTPVSGNLKVEFYNSNPSDTT 348
732 SDDCGTTWVLIND-----VOYIGTORD 773
                                                                                           827 SPSPSPSPSRSPSPSASPSPSSSPSPSSSPSSSPSTPSSSPVSGGVKVQYKNNDSAPGD 886
                                                                                                                                                                                                                                                                                          887 NQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAI------GCGNI 937
                                                                                                                                                                                                                                                                                                                      Length 493;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 208.5; DB 4;
Pred. No. 3.6e-06;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Kongsbak, Lars
TITLE OF INVENTION: No. 6187580el Pectate Lyases
FILE REFERENCE: 5378.200-US
CURRENT APPLICATION NUMBER: US/09/198,955A
CURRENT FILING DATE: 1998-11-24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSEQ for Windows Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION NUMBER: 1343/97
PRIOR PILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 1344/97
PRIOR FILING DATE: 1997-11-24
PRIOR PELICATION NUMBER: 60/067, 249
PRIOR FILING DATE: 1997-12-02
PRIOR FILING DATE: 1997-12-02
PRIOR FILING DATE: 1997-12-02
PRIOR FILING DATE: 1998-05-06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09/184,217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  :09-198-955A-12
Sequence 12, Application US/09198955A
Patent No. 6187580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 thermocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: Schulein, Martin
APPLICANT: Lange, Niels E.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Moller, Soren
APPLICANT: Glad, Saneo O. S.;
APPLICANT: Kaupphnen, Markus S.
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Kirk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4.18; 23.58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
```

19;

ò

ò 염 ò q δ g

ò g ò g ò g g

ò

```
530 ---TLPPLSAYHIVLKADSTEPVNSDLVVQYKDGDRNNATDNQIKPHFNIQNKGTSPVDL 586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        733 -----DDCGTTWVLINDDQHQYGN-WGQAITGDHANLRRVYIGTNGRGIVYGDIGGAPSG 786
                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                         542 PSSQPN-----DRHVAFSTDGGK-----NWFQGSEPGG--VTTGG----TVA 577
                                                                                                                                                                                                                                                                                                                                                                                                        275 PSLQGNYSWFIDYYLDQMKNAHTQNGKRLLDVLDVHWYPEAQGGGQRIVFGGAGNIDTQK 334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                432 KLYRNYDGNKS------GFGSIKVDAATSDTENSSVYASVTDEENSELHLIVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                       578 ASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           638 RSTDGGVT-----FQPVAAGLPSSGAVGVMFHAVPGKEG----DLW------LAASS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            680 GLYHSTNGGSSWSAITGVSSAVNVGFCKSAPGSSYPAVFVVGTIGGVTGAYRS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                847 SSSPSPSSSPSSPSPTPSSSPVSGGVKVQYKNND-SAPGDNQIKPGLQVVNTGSSSVDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            479 NKNFDDPINATFOLSGDKTYTSGRVWGFDQTGSDITEQAAITNINNNOFTY-----
                                                                                                                                                                                                                                                   Query Match 3.9%; Score 200.5; DB 2; Length 700; Best Local Similarity 21.7%; Pred. No. 1.9e-05; Matches 102;. Conservative 52; Mismatches 150; Indels 167;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               906 STVTVRYWFTRDGGSSTLVYNCDWAAIGCGNIRASFGSVNPATPTADTYLQ 956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GENERAL INFORMATION:
APPLICANT: Bjornvad, Mads E.
APPLICANT: Hatekeyama, Mariko
APPLICANT: Ghulelin, Martin
APPLICANT: Schulelin, Martin
APPLICANT: Schulelin, Martin
APPLICANT: Schulelin, Martin
APPLICANT: Gholelin, Martin
APPLICANT: Melsen, Jack B.
TITLE OF INVENTION: Endo-Beta-1, 4-Glucanases From
TITLE OF INVENTION: Endo-Beta-1, 4-Glucanases
FILE REFERENCE: 5195.200-05
CURRENT APPLICATION NUMBER: US/09/109,841
CURRENT FILING DATE: 1998-07-02
EARLIER PELICATION NUMBER: 0812/97
EARLIER FILING DATE: 1997-07-04
EARLIER FILING DATE: 1997-07-11
EARLIER PILING DATE: 1997-07-23
NUMBER APPLICATION NUMBER: 60/053,506
EARLIER PAPLICATION NUMBER: 60/053,506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOFTWARE: FastSEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: PRT ORGANISM: Saccharothrix australiensis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 2, Application US/09109841 Patent No. 6207436
                                                 SEQUENCE CHARACTERISTICS:
LENGTH: 700 amino acids
TYPE: AMINO ACID
  TELEFAX: 212 867 0298 INFORMATION FOR SEQ ID NO:
                                                                                                                        ; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-862-588B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQ ID NO 2
LENGTH: 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 8
US-09-109-841-2
                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
  19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Joergensen, Per Linaa
APPLICANT: Sch lein, Martin
APPLICANT: Sch lein, Martin
APPLICANT: Hansen, Christian
TITLE OF SEQUENCES:
TORESPONDENCE ADDRESS:
ADDRESSE: No. 59167960 No. 5916796th America, Inc.
STREET: 405 Lexington Avenue, 62nd floor
CITY: New York
STATE: New York
CITY: New York
                                                                                                                                                                                                                                                                                                                                                                                              SDNYNRTITFHHNWFENLNSRVPSFRFGEGHIYNNYFNKIIDSGINSRMGARIRIENN-- 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     63 GTITTSNTSASKIDVKDVSNVSIVGSGTKGE----LKGIGIKIWRANNIIIRNLKIHE-- 116
                                                                                                                                                                                                    -----VASG--DKDAIG----IEGPSKNIWVDHNE-LY 142
                                                                                                                                                                                                                                                                                                     143 HSLNVDKDYYDGLFDVKRDAEYITFSWNYVHDGWKSMLMGSS-------D 185
                                                                                                                                                                                                                                                                                                                                                      SDDCGTTWVLIND-------DQHQYGNWGQAITGDHANLR---RVYIGTNGR 773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   623 VNPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLY 682
                                                                                                                                                                                                                                                   HSTNGGSSW-----SAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYR 731
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAI------GCGNI 937
                                                    568 GGVTTGGTVAASAD----GSRFVWAPGDPGQPVVYAVGFG-NSWAASQGVPANAQIRSDR 622
                                                                                                                                                                                                                                                                                                                                                                                                                                                     774 GIVYGDIGGAPSGSPSPSVS-PSASPS---LSPSP---SPSSSPSPSPSSSPSSPSSPSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----LFENAKDPIVSWYSSSPGYWHVSNNKFVNSRGSMPTTSTTTYNPPYSYSLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SPSPSPSRSPSRSPSSSPSSSSPSSSSPSPSSSPSSSPVSGGVKVQYKNNDSAPGD
  Indels 125;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PatentIn Release #1.0, Version #1.25
  Mismatches 149;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME: Zelson, Steve T. / Lambiris, Elias J. REGISTRATION NUMBER: 30,335 / 33,728
REPERENCE/DOCKET NUMBER: 3425.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DK 164/90
FILING DATE: 19-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/DK91/00013
FILING DATE: 18-JAN-1991
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               OF EAST-
SOFTWARE: Patenta....
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/862,588B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
BDIUM TYPE: Floppy disk
CMMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 2, Application US/07862588B Patent No. 5916796
GENERAL INFORMATION:
  62;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ::| :: :| ||||||:
408 KGTFVKMSSSTNNADTYLE 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RASFGSVNPATPTADTYLQ 956
  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY: U
ZIP: 10017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-07-862-588B-2
Matches 103;
                                                                                                                                                                                                    117
                                                                                                                                                                                                                                                                                                                                                                                                      981
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      827
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             938
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      244
```

ä

g

ð 셤 à g g δŽ g

Q ò

ŏ

ò

```
869 VSGGVKVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 616;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 412;
                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 18, Application US/08313288B
Patent No. 5750502
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION:
CLONING, EXPRESSION AND USES OF A
TITLE OF INVENTION:
NOVEL SECRETED PROTEIN, F-SPONDIN
NUMBER OF SEQUENCES:
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOCFWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,288B
FILING DATE: January 5, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
                                                                                                                                                                                                                                                                                                                                                            Query Match 3.9%; Score 199.5; DB 4; Best Local Similarity 44.3%; Pred. No. 1.9e-05; Matches 39; Conservative 14; Mismatches 34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 1;
  REFERENCE/DOCKET NUMBER: 1997US001/CIP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
TELECOMMULTOATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3.7%; Score 191.5;
                                                                                                                                                                                                STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 47:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             929 WAAIGCGNIRASFGSVNPATPTADTYLQ 956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         60 WAQIGASNVTFNFVKLSSGVSGADYYLE 87
                         TELECOMMUNICATION INFORMATION
                                            TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
                                                                                                                                  SEQUENCE CHARACTERISTICS:
LENGTH: 616 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO: 18: SEQUENCE CHARACTERISTICS:
                                                                                        TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 47:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 412 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; MOLECULE TYPE: peptide US-08-313-288B-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CITY: New York
STATE: New York
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-313-288B-18
                                                                                                                                                                                                                                                                                             US-09-136-574A-47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION: Compositions and Methods for
Treating Cellulose Containing Fabrics Using Truncated
Cellulase Enzyme Compositions
                                                                                                               15;
                                                                                                                                                                                                                                                                                                                                     685 INGGS-----SWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTT 738
                                                                                                                                                                                                                                                                                                                                                                                                                             739 WVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDIGGAPSGSPSSVSPSASP 798
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      799 SLSPSPSPSSSSPSPSPSSSPSSSPSPSPSPSPSPSPSRSPSPSASPSSSSSSPSSS 858
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SPSPTPSSSPVSGGVKVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDG 918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            568 GGVTT-GGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPK 626
                                                                                                                                                                                                                                               627 TFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEG--DLWLAASSGLYHS 684
                                                                                                               Gaps
                                                                                                                                                                                                   GGVTALTSNIAQAAAGCR----PL-VDYAV--TSQWPGGFGAAVTVTNLGD---PL 70
                                                                                                                                                                                                                                                                                           ------PDGGCVQQLW----NGV-HS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             90; Indels 163;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STREET: Spring House Corporate Center, P.O. Box 457 CITY: Spring House
                                                                                                                                                                                                                                                                                                                                                                                                                                                         123 ------GSWNGA---NNAPTSFTLNGTSCNGAVGG------
                                                              Score 199.5; DB 4
Pred. No. 1.4e-05;
                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: US/09/136,574A
                                                                                                          36; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: US 08/932,571 FILING DATE: September 19, 1997 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: Howson and Howson
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 47, Application US/09136574A Patent No. 6294366 GENERAL INFORMATION:
APPLICANT: Farrington, Graham P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE: 19-Aug-1998
CLASSIFICATION: <Unknown>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bergquist, Peter
Daniels, Roy
Morgan, Hugh W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Williams, Diane P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Anderson, Paige
Gibbs, Moreland
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
                                                              3.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRIOR APPLICATION DATA:
                                                                                                          77; Conservative
                                                                                                                                                                                                                                                                                         71 SSWELS-WTF-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY: USA
                                                                 Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ZIP: 19477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              233 SDKDLL 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     919 GSSTLV 924
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 9
US-09-136-574A-47
US-09-109-841-2
                                                                                                          Matches
                                                                                                                                                                                                   27
```

g

ö

ò 셤

8

ŏ

g

ò

```
472 YKVEFNTPDDQITTPYIVVVNGHIDPNSKGD--LALRSTLY-GYNSNIIWRSMSWDNEVA 528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         284 T------GFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRIS 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDL--TKWDSGGQIHIAPMV-- 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STAKVPPIMAG--DQVLANGVIDSD-----GNVIYTFTDYVNTKDDVKATLTMPAYIDP 341
                                                                                                                                                                                                                                      701 VNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDH 760
                                                                                                                                                                                                                                                                                                                                                 224 TDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAP 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----YFGYSGLTIDRQHPNTIMVATQISWWPDT 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  372 IIFRSTDGGATWTRIWDWTSYPNRS----LRYVLDI--SAEPWLTF------GVQ 414
        374 KFYNLSI---KGTIDQIDKTNNTYRQTIYVNPSGDNVI-----APVLT-GNLKPNTDSNA 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               164 LGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANP 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            184 PSINESAPQSTDASNKDVVNQAVNTSAPRMRAFSLAAVAADAPAAGT-DITNQL----T
                                                                   585 FVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNG----TFYRSTD
                                                                                                         425 LI----DQONTSIKVYKVDNAADLSE-----SYFVNPENFEDVTNSVNITFPNPNQ
                                                                                                                                                     642 GCVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSS-A
                                                                                                                                                                                                                                                                              529 FNNGSG-SGDGIDKPVV-----PEQPDEPGEIEPIPEDSDSDPG----SDSGSD
                                                                                                                                                                                                                                                                                                                           761 ANLRRUYIGTNGRGIVYGDIGGAPSGSPSPSVSPSASPSLSPSPSPSSSPSPSPSSSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SDTKTSSNTNNGETSVAQNPAQQETTQSSST---NATTEETPVTGEATTTTTNQANTPAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 2, Application US/09421868
Patent No. 6177084
GENERAL INFORMATION:
APPLICANT: FOSTER: Timothy J.
TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
TITLE OF INVENTION: 1999-10-19
CURRENT APPLICATION NUMBER: US/09/421,868
CURRENT FILING DATE: 1999-10-19
PRIOR FILING DATE: 1999-10-19
PRIOR FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 933;
                                                                                                                                                                                                                                                                                                                                                                                                                                    3.7%; Score 189; DB 4; L. 23.6%; Pred. No. 0.00016; Live 80; Mismatches 326;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P----VPSTDTAND-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: Staphylococcus aureus
US-09-421-868-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PatentIn Ver. 2.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches 181; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NUMBER PATE SOFTWARE: PATE SEQ ID NO 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SULT 12
-09-421-868-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    290
                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                  g
                                                                                                                                                   ò
                                                                                                                                                                                           QΩ
                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                   ò
                          7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDL--TKWDSGGQIHIAPMV-- 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                290 STAKVPPIMAG--DQVLANGVIDSD-----GNVIYTFTDYVNTKDDVKATLTMPAYIDP 341
                                                                                           ------EWSPCS 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               164 LGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANP 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAP 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         T------GFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRIS 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              P-----VPSTDTAND------YFGYSGLTIDRQHPNTIMVATQISWWPDT 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         184 PSNNESAPOSTDASNKDVVNQAVNTSAPRMRAFSLAAVAADAPAAGT-DITNQL----T 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   372 IIFRSTDGGATWTRIWDWTSYPNRS----LRYVLDI--SAEPWLTF-------GVQ 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           238 NVTVGIDSGT-----TVYPHQAGYVKLNYGFSVPNSAVKGDTFKITVPKELNLNGVT 289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             KGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAE-- 528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ENVKKT------VDYEKYG 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --LNPSIIVRAGSFDPSSQPND--RHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSR 584
                                                                   KVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA--A 931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22 LVGTLIGFGLLSSKEADASENSVTQSDSASNES--KSNDS---SSVSAAPK--TDDTNV 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 2, Application US/08293728D
Patent No. 6008341
GENERAL INFORMATION:
APPLICANT: FOSLET, Timothy J.
APPLICANT: MODEVIL, Damien L.
TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene FILE REFERENCE: 05344.105011
CURRENT APPLICATION NUMBER: US/08/293,728D
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2.2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3.7%; Score 189; DB 3; Length 933;
Similarity 23.6%; Pred. No. 0.00016;
31; Conservative 80; Mismatches 326; Indels 180;
    Pred. No. 3.7e-05; ; Mismatches 36;
                                                                                                                                                 840 PSASPSPSSSPSSSPSSSPSPTPSSSP-----
                                                                                                                                                                                                                                                                            319 K---NNNNEEPSDKHIEQYLKKIKNSIST----
Best Local Similarity 20.8%; Promatches 41; Conservative 57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-293-728-2
                                                                                                                                                                                                                                                                                                                      932 IGCGN---IRASFGSVN 945
                                                                                                                                                                                                                                                                                                                                               : ||| :| || | 351 VTCGNGIQVRIKPGSAN 367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    933
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Sim
Matches 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-293-728-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local
                                                                                                                                                                                           259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        415
                                                                784
                                                                                                         199
                                                                                                                                                                                                                                      874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                529
```

g

ò 셤 ö g

셤

δ

g

ð

g

ŏ

g δý

õ

37;

US-08-760-797A-1

9

Thu Jul 18 16:01:07 2002

g ò g ŏ CP οy 음 ò g ð

```
PA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-932-929B-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                   Dp
                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                       δλ
                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
471 KGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAE-- 528
                                              -----GNVTLA--TGIGSTTANKTVL-----VDYEKYG 373
                                                                                                                                               374 KFYNLSI----KGTIDQIDKTNNTYROTIYVNPSGDNVI-----APVLT-GNLKPNTDSNA 424
                                                                                                                                                                                                                                                    425 LI----DQQNTSIKVYKVDNAADLSE-----SYFVNPENFEDVTNSVNITFPNPNQ 471
                                                                                                                                                                                                                                                                                                     642 GGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSS-A 700
                                                                                                                                                                                                                                                                                                                                                  472 YKVEFNTPDDQITTPYIVVVNGHIDPNSKGD--LALRSTLY-GYNSNIIWRSMSWDNEVA 528
                                                                                                                                                                                                                                                                                                                                                                                                   701 VNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDH 760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      573 SN-----SDSGSDSGSDSTSDSGSDSASDSDSASDSDSASDSDSASDSDSASDSDSD 624
                                                                                                  --LNPSIIVRAGSFDPSSQPND--RHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSR
                                                                                                                                                                                                   585 FVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNG---TFYRSTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                     529 FNNGSG-SGDGIDKPVV-----DEQPDEPGEIEPIPEDSDSDPG----SDSGSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    821 SSSPSPSPSPSPSRSPSPSASPSPSSSSPSSSPSSSPTPSSS 867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 1, Application US/08760797A
Patent No. 5928902
GENERAL INFORMATION:
GENERAL THORMATION: Michel
APPLICANT: Cohen, Joseph
TITLE OF INVENTION: Hybrid Protein Between CS
TITLE OF INVENTION: from Plasmodium and HBsAG
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/760,797A
FILING DATE: 04-DEC-1996
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/442,612
FILING DATE: 17-MAY-1995
ATTORNEY AGENT INFORMATION:
NAME: Baumeister, Kirk
REGISTRATION NUMBER: 33,833
REFERENCY/COKET NUMBER: B45015-1C2
TELECHONE: 610-270-5096
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        INFORMATION FOR SEQ ID NO: 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         . 423 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEFAX: 610-270-5090
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskett
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                            342 ENVKKT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STATE: PA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-760-797A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGIH:
```

```
845 SPSSSPSPSSSPSSSPSTPSSSPVSGG------VKVQYKNNDSA 883
                                                                                                                                                                                                                               63 NANPNANPNANPNANPNANPNKNNQGNGQGHNMPNDPNRNVDENNANNAVK --- NNNNEE 119
                                                                                                                                                                                                                                                                                                        :|: :|: :|:
                                                                                          787 SPSPSVSPSASPSLSPSPSSSP--SPSPSSSPSSSPSPSPSPSPSPSPSRSPSPSASP 844
                                               56; Gaps
                                                                                                                         884 PGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA--AIGCGN---IR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3.6%; Score 185.5; DB 4; Length 424;
  Length 423;
                                               38; Indels
; Score 186; DB 2;
; Pred. No. 8.9e-05;
54; Mismatches 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: De Wilde, Michel
APPLICANT: Cohen, Joseph
TITLE OF INVENTION: Hybrid Protein Between CS
TITLE OF INVENTION: From Plasmodium and HBSAG
CORRESPONDENCE: 4
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
SURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/932,929B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE/DOCKET NUMBER: 33,833
REFERENCE/DOCKET NUMBER: B45015-1FWC2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5096
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE:
CLASSIFICATION
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/760,797
FILING DATE: 04-DEC-1996
APPLICATION NUMBER: 08/42,612
FILING DATE: 17-MA-1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                         120 PSDKHIEQYLKKIKNSIST------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 1, Application US/08932929B Patent No. 6169171
GENERAL INFORMATION:
  3.6%;
20.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS: LENGTH: 424 amino acids
                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME: Baumeister, Kirk
REGISTRATION NUMBER: 3
  Query Match
Best Local Similarity
Matches 39; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; TYPE: amino acid
; STRANDEDNESS: Sir;
; TOPOLOGY: linear
US-08-932-929B-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER: IBM COI
                                                                                                                                                                                                                                                                                                                                                                      939 ASFGSVN 945
                                                                                                                                                                                                                                                                                                                                                                                                                 155 IKPGSAN 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              USA
```

```
Search completed: July
Job time: 54 sec
                                                                                                                        604 NSWAA----
                                                                                                                                                         q
                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                     ŏ
                                                                                                                        δ
                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               437 NSDRMLYGTGATL---YATNDLTKWDSGQIHIAPMVKGLEETAVNDLISPPSG--APLI 491
                                                                                                                                                492 SALGDLGGFTHADVTAVPSTIF----TSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQPN 547
                                                  SPSPSVSPSASPSLSPSPSPSSSP--SPSPSSSPSSSPSPSPSPSPSPSRSPSPSASPSASP 844
                                                                                                                        ----VKVQYKNNDS 882
                                                                                                                                                                                           883 APGDNQIKPGLQVVNTGSSSVDLSTVTVRYMFTRDGGSSTLVYNCDWA--AIGCGN---I 937
                                                                                                                                                                                                                                --- EWSPCSVTCGNGIQV 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels 123; Gaps
                                                                      Length 1719;
                                                                                                                                                                                                                                                                                                                                                                                        Sequence 4, Application US/08459568
Patent No. 5811304
GENERAL INFORMATION:
APPLICANT: Huang, Shi
TITLE OF INVENTION: Retinoblastoma Protein - Interacting
TITLE OF INVENTION: Zinc Finger Proteins
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
FILLING DATE: US/08/459,568
FILLING DATE: 02-JUN-1995
 Pred. No. 9.7e-05;
1; Mismatches 38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3.6%; Score 185.5; DB 2;
24.1%; Pred. No. 0.00061;
tive 55; Mismatches 199;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,411
FILING DATE: 06-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/POCKET NUMBER: 9-LJ 1264
TELECOMMUNICATION INFORMATION:
                                                                                                                      845 SPSSSPSSSSPSSSPSSSPVSGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
Best Local Similarity 20.7%; Pr
Matches 39; Conservative 54;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SED ID NO: 4:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1719 amino acids
                                                                                                                                                                                                                   Best Local Similarity 24.18
Matches 120; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CITY: San Diego
STATE: California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            amino acid
                                                                                                                                                                                                                                                                 938 RASFGSVN 945
                                                                                                                                                                                                                                                                                                   155 RIKPGSAN 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE T
US-08-459-568-4
                                                                                                                                                                                                                                                                                                                                                                     US-08-459-568-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                     셤
                                                                                                                        ö
                                                                                                                                                         g
                                                                                                                                                                                           ò
                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ô
```

```
961 QLPPLLIPT-----DPSSPPPCPPVLTVATPPPPLLPTVPLPAPSSSASPHPCPSPL 1012
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----SPQHSPA 750
                                                                      548 DRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFVWAPGDPGQ----PVVYAVGFG 603
                                                                                                                                        751 LRDFGKPSDGKAAW---TDAGLTSKKSKLESHSDSPAWSLSGRDERETVSPPCFDEYKMS 807
                                                                                                                                                                                                                                                                      808 KEWTASSAFSSVCNQQPLDLSSGVKQKAEGTGKTPVQWESVLDLSVHKKHCSDSECKEFK 867
                                                                                                                                                                                                                                                                                                                                                       647 ----QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAV 701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            702 NVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHA 761
                                                                                                                                                                                                                                                                                                                                                                                                                      868 ESHSVQPTCSAVKKRKPTTCMLQKVLLNE---YNGIDLPVENPADGTRSPSPCKSLEAQP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         762 NLRRVYIGINGRGIVYGDIGGAPSGSPSPSVSPSASPSLS----PSPSSSSPSPSPSPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     818 S-----SPSSSPSPSPSPSPSP---SRSPSP---SASPSPSSSPSSSPSS
                                                                                                                                                                                                              ----SQGVPANAQ-IRSDRVNPKTFYALSNGTFYRSTDGGVTF-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2, 2002, 09:12:46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1073 SPSPPPLSAISSVVSSG 1089
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              859 SPSPTP---SSSPVSGG 872
```

.

á

prote

Perfect score:

Sequence:

Scoring table:

Searched:

Minimum DB Maximum DB

Database

Result No.

```
A;Reference number: 221572
A;Accession: T35237
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Holecule type: DNA
A;Residues: 1-890 <SEE>
A;Residues: 1-890 <SEE>
A;Cross-references: EMBL:AL031515; PIDN:CAA20642.1; GSPDB:GN00070; SCOEDB:SC5C7.30c
                                                                                                  proteophosphoglyca
proline-rich prote
hypothetical prote
proline-rich prote
mucin, submaxillar
hydroxyproline-ric
spore germination
hypothetical prote
proline-rich prote
                                 aggrecan precursor
cell wall surface
UL36 protein - hum
                 mucin MUC5B, trach
                                                                                                                                                                                                                                                                                                                                                                                                                  C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C:Accession: T35.27 #
R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    69 VFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVW 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILY 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSS-SS 247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEAMAIDPFNSDRMLYGTGATLYATNDLTKW-DSGQIHIAPMVKGLEETAVNDLISPPS 486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11 MRSRRLVSLLAATASFAVAAALGVLPIAITASPA--HAATTQPYTWSNVAIGGGGFVDGI 68
                                                                                  hypothetical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          protein TPX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 MRRTRILIVILIALAAGLLA---GSPPAASAAEPAPRAAVAADSYTWKNARIDGGGFVPGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels 126;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 890;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 48.6%; Pred. No. 1.8e-102; Matches 475; Conservative 129; Mismatches 248;
                                                                                                                                                                                                                                                                                                                                                                                                   probable secreted cellulase - Streptomyces coelicolor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            47.4%; Score 2436; DB 2; 48.6%; Pred. No. 1.8e-102;
                                                E95206
WMBEH6
AC2284
T46707
T16586
T16586
T10599
J02220
B35621
T13369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Experimental source: strain A3(2) C; Genetics:
 360
33570
24132
24133
241
3449
13055
1308
13288
13288
13288
13288
13288
13288
13288
13288
13288
13288
13288
13288
13288
13288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C;Genetics:
A;Gene: SCOEDB:SC5C7.30c
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    368
 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    öγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Qγ
                                                                                                              2, 2002, 09:11:52; Search time 54.67 Seconds (without alignments) 1682.047 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sialidase - Actino
mannan endo 1,4-be
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            hypothetical prote protein TPX-VT3 -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S-layer protein -
proline-rich myros
proline-rich prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                probably secreted
Avicelase III - As
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    cellulase (EC 3.2. hypothetical prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            hypothetical prote cipA protein - Clo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                hypothetical prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  xylanase - Caldice
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1,4-beta-glucanase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              exò-alpha-sialidas
cellulase (EC 3.2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  endoglucanase I (E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        proline/lysine-ric
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      hypothetical prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            cellulase (EC 3.2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      proteoglycan core
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                probable secreted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Description
                                                                                                                                                                                                      5135
1 MDRSENIRLTMRSRRLVSLL.....RASFGSVNPATPTADTYLQX
                 GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd
                                                                                                                                                                                                                                                                                                                                              hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                            283138 segs, 96089334 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                  OM protein - protein search, using sw model
                                                                                                                                                                                                                                                        BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              D97013
T00349
F72393
T31085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       S02711
T34433
S15921
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A83412
S36859
S20590
A43802
A47704
T18262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  T31337
T17120
S50755
A49227
A48954
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S75111
S75138
T10361
S50754
T34434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                T35237
                                                                                                                                                                                                                                                                                                                                                                              length: 0
length: 2000000000
                                                                                                                                                                                        US-09-917-376-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                pirl:*
pir2:*
pir3:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PIR_71:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query
                                                                                                                                                                                                                                                                                                                                                                                                                                  Post-processing:
                                                                                                                                                                                                                                                                                                                                              Fotal number of
                                                                                                                                                                                                                                                                                                                                                                                sed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2436
10037
1161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11161
11
```

19;

Gaps

297 367 427

~

```
Avicelase III - Aspergillus aculeatus

Avicelase III - Aspergillus aculeatus

C; Species: Aspergillus aculeatus

C; Species: Aspergillus aculeatus

C; Accession: T00349

R; Arai, m; Taxada, G; Kawaguchi, T; Sumitani, J.

Submitted to the EMBL Data Library, June 1998

A; Reference number: Z14141

A; Reference number: Z14141

A; Accession: T00349

A; Reference number: Z14141

A; Accession: T00349

A; Reference number: Z14141

A; Accession: T00349

A; Residues: Preliminary; translated from GB/EMBL/DDBJ

A; Residues: 1-856 < ARA>

A; Residues: 1-856 < ARA>

A; Cross-references: EMBL; AB015511; NID: d1199887; PID: d1029971
  186
                                                                                                                                                                                                                                                                                                                                  346
                                                                                                                                                                                                                                                                                                                                                                                                          466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                 545 QPNDRHVAFSTDGGKNWFQ-GSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VWAAVGMYTNSWDPNDGAILRSSDQGATWOITPLPFKLGGNMPGRGMGERLAVDPNNDNI
                                                                                                                        247 SLGQASKTIFVGVAD-PNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATS
                                                                                                                                                                                                                      NTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQI
                                                                                                                                                                                                                                                                                                                                                                               366 SWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFG-VQPNPPVPSPKL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEETAVNDLISP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF-QPVAAGLPSSGAVGVM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     485 PSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           663 FHAVPGKEGDLWLA-ASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         40;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 856;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ugal cellulose-binding domain homology from fungal cellulose-binding domain homology <FCB>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 33.9%; Score 1740; DB 2; Best Local Similarity 43.5%; Pred. No. 3.5e-71; Matches 366; Conservative 139; Mismatches 297;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GAPSGSPSPSVS 793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1: :|:||:
764 GS-QPTPTPSVT 774
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             fungal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Gene: aviIII
C;Superfamily: fun
F;823-854/Domain:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   m
                                                                                             187
                                                                                                                                                                                                                                     231
                                                                                                                                                                                                                                                                                   306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               782
  127
                                            g
                                                                                                                                            g
                                                                                                                                                                                                                                  Db
                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                             ò
                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      probably secreted stalidase, several ASP-boxes and dockerin domain [imported] - Clostrid C; Species: Clostridium acetobutylicum (C; Species: Clostridium acetobutylicum (C; Species: 14-Sep-2001 #text_change 14-Sep-2001 | Species: 16-Sep-2001 | Species: 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Molecule type: DNA
A; Residues: 1-839 < KUR>
A; Cross-references: GB:AE001437; PIDN:AAK78895.1; PID:915023820; GSPDB:GN00168
A; Experimental source: Clostridium acetobutylicum ATCC824
C; Genetics:
A; Gene: CAC0919
                                                                                                                                                                                                                                                                                                            645
                                                                                                                                                                                                                                                                                                                                                                                                                                                   780
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            840
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           782
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       900
547 NDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSW 606
                                                                                                                                                                                                                 587
                                                                                                                                                                                                                                                             AASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVAA-GLPSSGAVGVMFHA 665
                                                                                                                                                                                                                                                                                                                                                        VPGKEGDLWLAASS-----GLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVV 720
                                                                                                                                                                                                                                                                                                                                                                                                  646 LPGGEGDVWLAGGAADGPYGLWHSTDGGGTFTRLPGVDAADTVGFGKAAPGASYQTLFTS 705
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               811
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNK 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SRRLVSLLAA-----TASFAVAAALGVLPIAITASPAHAAT-TQPYTWSNVAIGGGGFVD 66
                                                                                                                                                                                           ---HIAPSTDNGANWFGGTDPSGVSGGGTVAAGADGSREVWSP--EGAGVQYTTGFGTSW
                                                                                             841 SASPSPSSSPSSSPSSSPSSPSPTPSSSPVSGGVKVQYKNNDSAPGDNQIKPGLQVVNTGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  901 S----SVDLS----TVTVRYWFTR--DGGSSTLVYNCDW-----AAIGCGNIRA---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NKRIVSMVAGLSIIFTTGF------VTHISAAANKAAASVQSQGYKWDNAKIGAGGYVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQP
                                                                                                                                                                                                                                                                                                                                                                                                                                                 GTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----TSDTGGGTDPGPGPDPTP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------ADVRLTNTGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 839;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 39.7%; Score 2037; DB 2; Length 83:
Best Local Similarity 49.5%; Pred. No. 1.5e-84;
Matches 392; Conservative 118; Mismatches 252; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----SFGSVNPATPTA 951
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           415
                                                                                                                 475
                                                                                                                                                                                                                                                           607
                                                                                                                                                                                                                                                                                                                                                                                                                                                 721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 812
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                940
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            55
                                                                    487
                                                                                                                                                                                                                 533
                                                                                                                                                                                                                                                                                                                                                      999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      783
                      셤
                                                                    ò
                                                                                                                 d
                                                                                                                                                           ογ
                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

18;

1

```
A; Accession: F72393
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-707 < ARN>
A; Cross-references: GB: AE001712; GB: AE000512; NID: 94980799; PIDN: AAD35393.1; PID: 9498
A; Cross-references: strain MSB8
C; Genetics: A; Geneti
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 5
T31085
Xylanase - Caldicellulosiruptor sp.
C;Species: Caldicellulosiruptor sp.
C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
A; Reference number: A72200; MUID:99287316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      617
                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ΩÝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ρ
                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C; Species: Thermotoga maritima
C; Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C; Accession: F72393
R; Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       seo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Nature 399, 323-329, 1999
A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DYAASTSTGTGAVALSADGDTVLLMSSTSGALVSKSQG---TLTAVSSLPSGAVIASDKS 583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       742
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    760
                                                                                                           145
                                                                                                                                                   119
                                                                                                                                                                                               202
                                                                                                                                                                                                                   265
                                                                                                                                                                                                                                                                                                                                                                                                                                                       443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              324
                                                                                                                                                                                                                                                                                                                                                                                                                297
                                                               63
                                                                                                                                                                                                                                                                                                                                                                                                TWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NDDQHQYGNWGQA--ITGDHANLRRVYIGTNGRGIVYGDIGGAPSG-----SPSPSV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VFWSRDGGATWQAVPGAPT-GFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             564 GSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GTGATLYATNDLTKWDSGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STNGGSSWSAI - TGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLI
                                        SSLALLCAALLGKLADAAASQAYTWKNVVTGGGGGFTPGIVFNPSAKGVAYARTDIGGAY
                                                                                                        RWDAANGRWIPLLDWVG----WNNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAI
                                                                                                                                    LRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RIWDWTSYPNRSLRYVLDISAEPWLFFGVQPNP-PVPSPKLGWMDEAMAIDPFNSDRMLY
                    AALGVLPIAITASPAHAATTQPYTWSNVAI-GGGGFVDGIVFNEGAPGILYVRTDIGGMY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        - Thermotoga maritima (strain MSB8)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      hypothetical protein TM0305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        l:
SA 822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SS 854
                                                                                                                                                                                                                                                                                                                                                                                                              238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         355
                                                                                                                                                                                                                                     120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           412
                    30
                                                                                                        89
                                                                                                                                                 64
                                                                                                                                                                                            146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         743
                                                                                                                                                                                                                                                                                 206
                                                                                                                                                                                                                                                                                                                                                                     266
                                                                                                                                                                                                                                                                                                                                                                                                                                                         325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               853
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        821
                                                             셤
                                                                                                                                               g
                                                                                                                                                                                          δ
                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                    ò
                                                                                                        à
```

```
27;
                                                                                                                                                                                                                                                                                208
                                                                                                                                                                                                                                                                                                          328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             670
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         729
                                                                                                                                                                                                                                  119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       654
                                                                                                                                                                                      92 AANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQ 151
                                                   Gaps
                                                                                                                         278 YDVTPMKGD-----FGYCGIDVQE---NVVIVSTLDRWYPHDEIFISLNGGETWRPLLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  329 KANF------DINKAPWIK---DLNP------HWISD-VKIDPEDMNRAIFTTGYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            536 FPASDPVNPSKFYIFDWKNGDFLISKDGGKSFMK-GAKLPSFDNWWVSLYSFPVLAPDRE
                                                                                                                                                                                                                                                                                                                                                                       209 OMTNFPDVGTYIANPIDTIGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFW
                                                                                                                                                                                                                                                                                                                                                                                                       SRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 449 LYATNDLTKWDSGGQIHIAPMVK-----GLEETAVNDLISPPSGAPLISALGDLGGFTH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         503 ADVTAVPSTIFTSPVFTTGTSVDŸAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     425 ESLDTPPSSMY-KPLKWTSLGIAFAYQNSKFVARVHTY---TYP---FLSYSEDGGINWR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q-GSEPGGVTTGG--TVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPA---NA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               478 EIETVPEGITDGGRLSLAVSNDGKTLVWSPAN--HEVIVSSDKGKSWKKAISVPVPEFNY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         671 GDLWLAAS-SGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      32 LGVLPIAITASPAHAATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWD
                                                                                                                                                                                                                                  61 EETKRWKQLFDFLRRDQSDYMGVLSVALDPSDPKRIYAMTGKYTQDW-AGYGAILISEDY
                                                                                                                                                                                                                                                                                152 GATWOITPLP---FKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        389 WISYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QIRSDRVNPKTFYAL - - SNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHA - - - - VPGKE
                                                98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      730 YRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDI 780
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 707;
Query Match 22.6%; Score 1161; DB 2; Length 70 Best Local Similarity 34.6%; Pred. No. 3e-45; Matches 267; Conservative 126; Mismatches 280; Indels
```

sp. Rt69B	s 53;	20	65 19	25 50	77	23 56	81 08	07 58	26	35 88	81 42	39 96	63 56	. 08	25 67
December 1997 Lase genes from Caldicellulosiruptor Dm GB/EMBL/DDBJ LD:g2760909; PIDN:AAB95	re 431.5; DB 2; Length 1779; ed. No. 5.7e-12; Mismatches 382; Indels 477; Ga	-RTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVSIAAD 120	DGAILRSSDQGATWQIT	GNMPGRGMGERLAVDPNNDNILYFGAPSGRGLWRSTDSGATWSQNTNFPDVGTYIANPTD 225 : : : : : : : : : : :	TTGYQSDIQGVVWVAFDKSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQ 277 :::	AVPGAPTGFIPHKGVFDPVNHVLXIATSNTGGPYDGSSGDVWKFSV 323	-TSGTWTRISPVPSTDTAND-YFGYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGA 381 	TWTRIWDWTSYPNRSLRYVLDISAEP	ASDTEKKWVLKHFNSITAGNEMKPSELLISENNYNFSKADEFVNFATSNNIAIRGHTLVW 428	MDEAMAIDP 435 :: HEQTPDWFFKDANGNTLSKDALLSRLKQYIYTVVGRYKGKVYAWDVVNEAIDESQGNGFR 488	FNSDRMLYGTGATLYATNDLTKWDSGGOIHIAPMVKGLEETAVNDL 481 :	ISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGS 539	FDPSSQPNDSTDGGKNWFQ 563 	GS	DRVNP 62 KIENVGTTAVDLSRVKV 76
Dmitted to the EMBL Data Library, Description: Family 10 and 11 xyl. Reference number: 220972 Accession: T31085 Status: preliminary; translated f. Molecule type: DNA Residues: 1-1779 < MOR> Cross-references: EMBL:AF036924; I Genetics: Note: xynC	8. imilarity 19. ; Conservative	VDGIVFNEGAPGILYV :: :: GTLIFHQEAKAAAYTVDF	PINTNKVWAAVGMYTNSWDPN : : :NRSSIWDGVAV			AVPGAPTGFIPHKGVFDP 	_					IS			DGSREVWAPGDPGQPV
A. Reference numb A. Reference numb A. A. Accession: T31 A. Status: prelim A. Wolecule type: A. Residues: 1-17 A. Cross referenc C. Genetics: A. Note: XynC	Query Best Matc}	Qy 65 Db 20	Qy 121 Db 72	Qy 166 Db 120	Qy 226 Db 151	Oy 278 Db 202	Qy 324 Db 257	Qy 382 Db 309	Qy 408 Db 369	Oy 427 Db 429	Oy 436 Db 489	Qy 482 Db 543	Qy 540 Db 597	Qy 564 Db 657	Qy 581 Db 717

```
1,4-beta-glucanase (EC 3.2.1.-) - Anaerocellum thermophilum (fragment)
C;Species: Anaerocellum thermophilum
C;Species: Anaerocellum thermophilum
C;Species: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C;Accession: T31337
R;Zverlov, V.; Mahr, S.; Riedel, K.; Bronnenmeier, K.
Microbiology 144, 457-465, 1998
Microbiology 144, 457-465, 1998
Omains.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      40;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A.Reference number: 221003; MUID:98154434
A.Roccession: T3137
A.Status: preliminary; translated from GB/EMBL/DDBJ
A.Status: preliminary; translated from GB/EMBL/DDBJ
A.Molecule type: DNA
A.Residues: 1-1711 <2VE>
A.Cross-references: EMBL:286105; NID:e1071329; PID:e350354; PIDN:CAB06786.1
C.Genetics:
A.Gene: celA
C.Keywords: glycosidase; hydrolase
                                                                                                                                                                                                                                                                                                                                                                                                                 947 VVKLGANAGGADYYVEVGFKSGAGVLAAGQSTKEIRLSIQKSSGSYNQSNDYSVRSANSY 1006
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | : : | : : | : : | | 887 NGLSSPINVLNPKIKIENVGTTAVDLSRVKVRYWYTIDGEATQSVSVASSINPAYIDVR 946
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        829 SPSPSPSRSPSPSASPSPSSSPSPSSCSPSSS----SPSPTPSSSPVSG-GVKVQYKNNDSA 883
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  884 PGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAIGCGNIRASFGS 943
626 KTFYALSNGTFYRSTDGGVTFQP------VAAGLPSSGA-------VGVMFHAVP 667
                                                                                                                                                                                                                                                                                          765 RVYIGTNGRGIVY----GDIGGA----- 787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      316 GDVWKFS---VTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQI----- 365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 366 SWW-----PDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQP 415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 416 NPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEE 475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels 281; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      262 PNNPVFWSRDGGATWQA----VPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDGS--S 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    74 WAVYEYKDAFVKSGQLEHI--LNQIEWVNDYF-----VKCHPSKYVYYYQVGDGSKDH 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25 PN-----WVRN----NWRGDSALKDGQDNGLDLTGGWFDAGDHVKF----NLPMSYTGTMLS 73
                                720 VGTIGGVTGA----YRSDDCGTT------WVLINDDQHQYGNWGQAITGDHANLR
                                                                                             668 GKEGDLWLAASSGLYHSTNGGSSWSA -----ITGVSSAVNVGFGKSAPGSSYPAVFV
                                                                                                                                                                                                                                                                                                                                                                                          Length 1711;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 7.8%; Score 398.5; DB 2; Best Local Similarity 22.9%; Pred. No. 1.7e-10; Matches 197; Conservative 116; Mismatches 267;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    : || ||:
1184 LTSGVSGADYYLE 1196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     944 VNPATPTADTYLQ 956
                                                                                                ò
                                                                                                                                        Dp
                                                                                                                                                                                          ò
                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Db
                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

C;Keywords: glycosidase; hydrolase	Query Match 7.7%; Score 397.5; DB 2; Length 1742; Best Local Similarity 21.7%; Pred. No. 1.9e-10; Matches 189; Conservative 109; Mismatches 269; Indels 303; Gaps 35; Qy 262 PNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLY-IATSNTGGPYDGSSG 316	317 DVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIWVATQISWWSWW	369PDTIIFRSTDGGATWTRIWDWTSYPNRSLRXVLDISAEPWLTFGYQPNPPV 152 PAEVMQMERPSFKVTQSSPGSA	Qy 420 PSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEETAVN 479	QY 480 DLISPPSGAPLISALGDLGGFTHADVTAVRSTIFTSPVFTTGTSVDY 526 Dh 225 GYYNGWGFFVDFI.SWAAVWIVI.ATHINGTYLFTRARSYONWWAYTGGSNTTDYKWAH 279	527GKNW 527GKNW 5280 CWDDVHNGAALLARITDKDTYKQIIESHLDXWTTGYNGERIKYFPKGLAWLDOWGS	QY 562 FQGSEPGGVTTGGTVAASADG-SRFVWAPGDPQQPVYAVGEG 603 Db 337 I.RYATTTAFLAFVYSDAGCPPGKETYRKFGFSOIDYALGSTGRSFVVGFGTNPPKR 394	604NSWAASQGVPA	OY 622RVNPKTFYALSNGTFYRSTDGGVTFOPVAAGLPSSGAVGVMFHAVPG 668 Dh 455 AlakmyllyggnDjDpFKATFFPPTNDFFFVERGTNASGTNFFFFRAT 501	669 KEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVF	Qy 719 VVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRR 765 :	OY 766 VYIGTNGRGIVYGDIGGAPSGSPSPSVSPSASPSLSPSPSPS 807 1 1 1 1 1 1 1 1 1	QY 808 SSPSPSPSSSPSSSPSPSPSPSPSPSASPSASPSPSSSPSSSPSSSPSSSPSSS 867 (111111111111111111111111111111111111	QY 868 P-VSGGVKVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYN 926 : : : : Db 701 PATSGQIKVLYANKETNSTINTIRPWLKVVNSGSSSIDLSRVTIRYWYTVDGERAQSAIS 760	Qy 927 CDWAAIGCGNIRASFGSVNPATPTADTYLQ 956 	RESULT 8 S50755 hypothetical protein VSP-3 - Chlamydomonas reinhardti1 C.Specias: Chlamydomonas reinhardti1 C.Specias: Chlamydomonas reinhardti1 C.Specias: Chlamydomones revision 21-Jul-1995 #text change 21-Jul-2000
Db 170 NPTKAATYLQHAKEY 197	QY 476 TAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY 526 Db 198 TAANGYYNSWSGFYDELSNAAVWLYLATNDSTYLTKAESYVQNWPKISGSNTIDY 252 QY 527GKNWFQ 563 QY 527GKNWFQ 563 Db 253 KWAHCWDDVHNGAALLLAKITGKDIYKQIIESHLDYWITGYNGERIKYTPKGLAWLD 309	564GSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFG	604NSWAASQGVPA	622RVNPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFH	OY 665 AVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAP 710 1	711GSSYPAVEVVGTIGGVTGAYRSDDGGTTWVLINDDQHQYGNWGQAITGDHANLRRV	QY 767 YIGTNGRGIVYGDIGGAPSGSPSPSVSPSASPSLSPSPSPSSSPSPSPSP 816	QY 817 SSSPSSPSPSPSPSPSPSPSPSPSPSPSSSPSSSPSSS	QY 876 QYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAIGCG 935 :-	Qy 936 NIRASFGSVNPATPTADTYLQ 956 : : : : Db 750 NVTFKFVKLSSSVSGADYYLE 770	T17120 T17120 (EC 3.2.1) precursor, thermoactive - Caldocellum saccharolyticum	C;becaes: Caludocallum SacchafolyLicum C;bate: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Nov-2000 C;Accession: T17120; A43745 R;Te'o, V.S.; Saul, D.J.; Bergquist, P.L.	Appl: microbiol: bloteching. 43, 291, 299. A;Title: CelA, another gene coding for a multidomain cellulase from the extreme thermoph A;Reference number: 218698; MUID:95336703 A;Accession: T17120 A;Cetatius: oraliminary: translated from CP,CMDE,CMDE,CMDE	Notecule type: DNA ;Residues: 1-1742 <teo> :Cross references: EMBL:L32742; NID:g537499; PID:g537500; PIDN:AAA91086.1</teo>	Appl. Environ. Microbiol. 57, 694-700, 1991 A; Lucelli, E.; Biana Jasmat, N.; Grayling, K.A.; Love, D.K.; Bergquist, P.L. A; Title: Cloning, sequence analysis, and expression in Escherichia coli of a gene coding A; Reference number: A43745; MUID:91247819 A; Accession: A43745	A;Status: preliminary A;Wosicuse type: DNA A;Resicuse: 1516-1544,'A',1546-1742 <lue> A;Cross:references: EMBL:M36063; NID:g144292; PIDN:AAA72860.1; PID:g144293 C;Genetics: A;Gene: celA</lue>

	OY 245 SSSLGOASKTIEVGVADPNNPVEWSRDGGATWQAVPGAPTGFIPHKGVFDPV 296 192SDKTINVAATESSFDDLGRQCHWGGLKPGKGAVYNCKPL 230 QY 297 NHVLXIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHP 356	Oy 417 -PPVPSPKLGWMDEAMAIDPENSDRMLYGTGATLYATNDLTKWDSGGQI 464	558 GKNWFOGSEPGCVTTGGTVAASADGSRFVWAP	1	Db 759 AAEPTQAPTVAPSVEPTQAPGAQPSSAPKPGATGRAPSVVNPKATGAAT 807 RESULT 10 A48954 mannan endo-1,4-beta-mannosidase (EC 3.2.1.78) - Caldocellum saccharolyticum N.Alternate names: beta-mannanase C;Species: Caldocellum saccharolyticum C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999 C,Accession: A48954; B4375 R;Gibbs, M.D; Saul, D.J; Luthi, E; Bergquist, P.L. Appl. Environ. Microbiol. 58, 3864-3867, 1992
C; Accession: S50755 R; Woessner, J.P.; Molendijk, A.J.; van Egmond, P.; Klis, F.M.; Goodenough, U.W.; Haring, Plant Mol. Biol. 26, 947-960, 1994 A; Title: Domain conservation in several volvocalean cell wall proteins. A; Reference number: S50754; MUID:95093034 A; Accession: S50755 A; Status: preliminary A; Molecule type: mRNA A; Residues: 1-473 < WOE> A; Cross-references: EMBL:L29029; NID:9530875; PIDN:AAB53953.1; PID:9530876	Ouery Match 1.3%; Score 372.5; DB 2; Length 473; Best Local Similarity 33.0%; Pred. No. 5.5e-10; Matches 132; Conservative 36; Mismatches 139; Indels 93; Ga 492 SALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSOPN 1	DB 85 AGADATVILSNRYYSYMDTEKMGSSPNKVRDLKDWNNAGGSLVLLDG 131 QY 602 FGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAV 659 132 YSTASGTNTFVQLIDAVLGTKAGSGCTGALTNGNVNYRRANSSSPFGKITSPLLVKGS- 190 QY 660 GVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPA 716 QY 660 GVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPA 716 Db 191SRGESGLTGCTSGAVLFSSNPDKMTKAVTASAITWSVGKGAITFIGSSFAM 241 QY 717 VFVYGTIGGVTGAXRSDDCGTTWVLINDDQHQYGNWGQAITGDHAULRVYIGTNGRGIV 776	Db 242 PHLKG	RESULT 9 A4927 stalidase - Actinomyces viscosus clobecies: Actinomyces viscosus C.Species: Actinomyces viscosus C.Species: Actinomyces viscosus C.Accession: A49227 R.Yeung, M.K. Infect. Immun. 61, 109-116, 1993 A.Title: Complete nucleotide sequence of the Actinomyces viscosus T14V sialidase gene: FA: Reference number: A49227; MUID:93114861 A.Contents: T14V A.Accession: A49227 A.Status: preliminary A.Molecule type: nucleic acid	A;Note: sequence extracted from NCBÍ backbone (NCBIN:121598, NCBIP:121599) Query Match Best Local Similarity 22.5%; Pred. No. 1.7e-08; Matches 227; Conservative 101; Mismatches 330; Indels 351; Gaps 48; Qy 30 AALGVLPIAITASPAHAATTQPYTWSNVAIGGGGFVDGIVFNEGAPGI-LYVRTD 83

j

```
Local Sim
les 142;
                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ω
                                                                                                                     A.Cross-references: GB:L01257; NID:g144290; PIDN:AAA71887.1; PID:g144291
A.Cross-references: GB:L01257; NID:g144290; PIDN:AAA71887.1; PID:g144291
B.Mote: sequence extracted from NCBI backbone (NCBIN:121576, NCBIP:121577)
B.Luethi, E.; Bhana Jasmat, N.; Grayling, R.A.; Love, D.R.; Bergquist, P.L.
Appl. Environ. Microblool. 57, 644-700, 1991
A.Title: Cloning, sequence analysis, and expression in Escherichia coli of a gene coding A.Reference number: A43745; MUID:91247819
A.Status: preliminary
A.Status: preliminary
A.Status: preliminary
A.Status: Treasumer: A43745 NID:9144292; PIDN:AAA72861.1; PID:g144294
A.Gross-references: EMBL:M36063; NID:9144292; PIDN:AAA72861.1; PID:g144294
A.Gross-references: EMBL:M36063; NID:g144292; PIDN:AAA72861.1; PID:g144294
A.Gross-references: EMBL:M36063; NID:g144294
A.Gross-references: EMBL:M36063; NID:g144294
A.Store: the authors translated the codon CAC for residue 262 as Glu
A.Note: this sequence has been revised in reference A48954
C.Keywords: glycosidase; hydrolase; polysaccharide degradation
a multidomain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           326 TPTPTSTVTPTPTPTPTPTVTATPTPTPTPTPTSTPATSGQIKVLY---ANKETNSTTNT 382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   383 IRPWLKVVNSGSSSIDLSRVTIRYWYTVDGERAQSAISDWAQIGASNVTFKFVKLSSSVS 442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       426 WMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEETAVNDLISPP 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----- 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 486 SGAPLISALGDLGGFTHADVTAVPS--TIFTSPVFTTGTSVDYAELNPSIIVRA---GSF 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----AIKALRD-AGFKHTIMVDAPNWGQDWSNTMRDNAQSIMEADPLRNLVFSIHMYGVY 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             234 NTASKVEEYIKSFVDKGLPLVIGEFGHQHTD-----------GDPDEE 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VV-----YAVG-FGNSWAASQGVPANAQIRS--DRVNPKTFYALSNGTFYRSTDGG--- 643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    271 AIVRYAKQYKIGLFSWSWCGNSSYVGYLDMVNNWDPNNPTPW-----GQWYKTNAIGTSS 325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ---WL----AASSGL-----YHSTNGGSSWSAITGVS--SAVNVGF----GKSAP 710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          494
A;Title: The beta-mannanase from "Caldocellum saccharolyticum" is part of A;Reference number: A48954; MUID:93119139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RVYIGTNGRGIVYGD----IGGAPSGSPSPSVSPSASPSLSPSPSSSPSPSPSPSSSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TATPTPTPTVTPTVTPTVTPTVTATP------TPTPTPTPTPVSTPATGGQIKVLYAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAIGCGNIRA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDD---QHQYGN---WGQAITGDHANLR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      541 DPSSQPNDRHVAFSTDG----GKNWFQGSEPGGVTTGGTVAASADGSRFVWAPGDPGQP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          443 GADY--YLEIGFKSGAGQLQPGKDTGEIQMRFNKDDWSNYNQGNDWSWIQSMTS----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----VTFQPVAAGLP----SSGAVGVMFHAVPGKEGDL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 1331;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   141 WKEIKSVLDG-NEDFVIINIGNEPYGNNNYONW------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 6.6%; Score 340.5; DB 2; Best Local Similarity 22.7%; Pred. No. 5e-08; Matches 140; Conservative 97; Mismatches 189;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---YGENEKVTAYIDGVLVWGQEPSGA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KFVKLSSSVSGADYYLE 652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SFGSVNPATPTADTYLQ 956
                                         A;Accession: A48954
A;Status: prellminary
A;Molecule type: nucleic acid
A;Residues: 1-131 <GIB>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                880
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          940
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
```

RESULT S76211

```
hypothetical protein slr0442 - Synechocystis sp. (strain PCC 6803)
C.Species: Synechocystis.sp.
A;Variety PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C;Accession: 576211
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas DNA Res. 3. 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys S.
A;Reference number: S74322; MUID:97061201
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Molecule type: DNA
A;Residues: 1-611 cKANA
A;Residues: 1-611 cKANA
A;Residues: 1-611 cKANA
A;Cross-references: EMBL:D90914; GB.AB001339; NID:g1653477; PIDN:BAA18470.1; PID:d101
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C;Accession: S75138
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas DNA Res. 3, 109-136, 1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DNA Res. 3, 109-136, 1996
A; Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            262 -EIRGNIVLAGCGSTDLTDAYIDDLNSIQNAAGGFAAVRASTPMPSIPLSALASARSQNR 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  798
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       836
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   429 EAMA--IDPFNSDR----MLYGTGATLYATNDLTKWDSGGQIHIAPWVKGLEETAVNDLI 482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ------VYTLNVGNNDSISLPRSGEQPVDGYYYYESSSLDRKFELNFDNTGGKK 368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----CMNGNSSRTTKAFV-LAPNYDIGRTGNGS-----YEGMLLGKT--YAKGNCAV- 454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ------NNGKLAVTQT-----SQTWGSVPPDL--RPLTAPSPTLTSFSSW 491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  739 WVLINDDQHQYGNWGQAITGDHANLRRYYIGTNGRGIVYGDIGGAPSGSPSPSVSPSASP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           628 FYALSNGTFYRSTDGGVT-----FQPV---AAGLPSSGAVGVMFHAVPGKEGDLWLAAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------WAASQGVPANAQIRSDRVNPKT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -DLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               207 TRLMVDIPISGKPGTSTFGGSDTVPALWISEGGVDKGTAGDTVDTTGTPAVPKTG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AELNPSIIVR-AGSFD-PSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hypothetical protein slr1753 - Synechocystis sp. (strain PCC 6803)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 611;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                837 SPSPSASPSSSPSSSPSSSPSSSPSPTPSSSPVS-----GGVKVQYKNN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6.0%; Score 308.5; DB 2;
26.7%; Pred. No. 5.6e-07;
live 51; Mismatches 180;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  369 YRLYVTGDIGRNAEINGTCNTADGCQPTDIQIFGLSSSGSI---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Reference number: S74322; MUID: 97061201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FVWAPGDPGQPVVYAVGFGNS------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    S-----PPSGAPLISALG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
```

œ

```
hypothetical protein 92 - Orgyla pseudotsugata nuclear polyhedrosis virus C;Species: Orgyla pseudotsugata nuclear polyhedrosis virus, OpMNPV C;Species: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000 C;Accession: Ti0361 R;Ahrens, C.A.; Russell, R.R.; Funk, C.J.; Evans, J.; Harwood, S.; Rohrmann, G.F. Yirology 229, 381-399, 1997 A;Tile: The sequence of the Orgyla pseudotsugata multinucleocapsid nuclear polyhedro A;Reference number: 217011; MUID:97271300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Cypecies: Chlamydomonas engametos
Cybecies: Chlamydomonas engametos
Cybecies: Chlamydomonas engametos
Cybecies: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 21-Jul-2000
Cybecession: 350754
Rywoessner, J.P.; Molendijk, A.J.; van Egmond, P.; Klis, F.M.; Goodenough, U.W.; Hari Plant Mol. Biol. 26, 947-960, 1994
A.Title: Domain conservation in several volvocalean cell wall proteins.
A.Reference number: 550754
A.Reference number: 550754
A.Reference number: 550754
A.Reference number: 150754
A.Residues: preliminary
A.Residues: 1-351 < WWD:
A.Residues: 1-351 < W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A. Residues: 1-279 CAHR>
A. Cross-references: EMBL:U75930; NID:g2934903; PIDN:AAC59091.1; PID:g1911338
C. Superfamily: proline-rich protein 3
                                                                                                                                                                                                                                                                                                                                                                        IGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRYYIGTNGR--GIVYGDI 780
     697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           78 SPIPTPSPILSPIPSPIPTPSPIPSPIPSPIPTPSPIPTPSPIPTPSPIPTPSPIPTPSPIPTP 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
     -SWSAITGV
                                                       DYLNTYGSOTSGDVYVEAPLDISIGSYIYTGGGGEPGNVFLQAGGDITTSYIDTSAANGG
                                                                                                                                                             DIFIQSGGDTEVGYLFTKGYEGRGGDVYVETGRYFRAIDGFLLGEEGPFSVYTAGLTVG-
                                                                                                                                                                                                                                                                     ------TNGTIGAISCOD
                                                                                                                                                                                                                                                                                                                       GGAPSGSP-----SPSVSPSASPSLSPSPSSSSPSPSSSSPS
                                                                                                           ---FGKSAPGSSYPAVFVVGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 279;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5.7%; Score 293; DB 2; L
55.8%; Pred. No. 1.1e-06;
iive 26; Mismatches 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5.7%; Score 291; DB 2; L
64.4%; Pred. No. 1.8e-06;
Live 10; Mismatches 18;
                                                                                                                                                                                                                                                                                                                                                                                                                             PSPSPSPSPSRSPSASPSSSSPSSSSPSSSPSSSPSSSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         hypothetical protein WP6 - Chlamydomonas eugametos
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10; Mismatches
                                                                                                     SSAVIVG
        ---SSGLYHSTNGGS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 843 SPSPSSSPSSSPSSSPSSSP 868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 64.48
Matches 58; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
        GDLWLAA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
Matches 48; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Accession: T10361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     14
                                                                                                                                                                                                                 723
                                                                                                                                                                                                                                                                                                                                                                                                                             826
                                                         844
                                                                                                                                                             904
                                                                                                                                                                                                                                                                                                                    781
     671
                                                                                                           869
                                                                                                                                                                                                                                                                     963
                                                    g
                                                                                                        ò
                                                                                                                                                          g
                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                               NID:g1652127; PIDN:BAA17052.1; PID:d101778
to the EMBL Data Library, June 1996
                                                                                                                                                                                                                                                                                              46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            79 QFGLDQQQIANFLSNPDIRNILTRIVGGDASIINGLIQVSGGNANLFLMNPAGMIFGPNA 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                254 MNQGGSFSVLDLPTLLTQGASNLDLG--LAVQPNGSVTTNGTNALVSPLPGSVTISGNVD 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             670
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          139 SINVPGDFVVTTGSAIGFGNDQWFQV----FSDNDYNALIGNPSQFAFDLANPGLIINA 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SGATWSQMTNFPDVGTYIANPTDTTGYQS---DIQ-----GVVWVAFD-KSSSSLGQASK 253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                254 TIFVGVADPNNPV----FWSRDGGA--TWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNT 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISW 367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      503 -GNLTFDSPTITIDALFS----LGTGSIIFANTGPINTGNTLVTSPFTNLDFDNKIQLNA 557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NTTFTAPGYDIYFRKSVNGGFDLLGNANFVYFDDGAGITTPLKSFGVTATEIYVGNDIVT 617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QGNQIFDGVFYGLQPVNLTSSAGSVIFTNNILLNGSLQVQTAQNIVSQPSSLSAVEIAS 677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LNAGGTLTTGY1ETSGTNGGDVTTSSGSNTSTAY1DTRGFGDGLEIDSLGGAVSIESKGD 785
                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                 78
                                                                                                                                                                                                                                                                                                                                                                           ::||| ::||;
LALLACSSSFSG----NVLAQNITPAPDGTGTTVDAQGNQFNIGGGSLSGDGQNLFHSLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             17 VSLLAATASFAVAAALGVLPIAITASPAHAATTQPYTWSNVAIGGG-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------GEVDGIVFNEGAPGILYVRTDIGGMYRWDA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ------IAADPINTNKVWAAVGMYTNSWDPNDGAILRSS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -DQGATWQITPLPFKLGGNMPGRGMGERLAVDPNND-----NILYFGAPSGKGLWRSTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  368 TLI----DSNSVVKADALLTGNGGTVIVWADDSTRFSGNI-----SAO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGTMGGNGG----FVETSG-------AKSLMVDDTARVNTFATMGELGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            368 W---PDTIIFRSTDGGATWTRIWD----WTSYPNRSLRYVLDISAEPWLTFGVQPNPPVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---TNDLTKWDSGGQIHI---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----APMVKGLEETAVN---DLISPPS-----GAPLISALGDLG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---TSVDYAELNP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIIVRAGSFDPSSQPNDRHVAF---STDGGKNWFQGSEPGGVTTG------GTVA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -- ASOGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKE--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match 5.9%; Score 301.5; DB 2; Length 1749; Best Local Similarity 19.0%; Pred. No. 3.9e-06; Matches 225; Conservative 117; Mismatches 402; Indels 439;
Status: nucleic acid sequence not shown; translation not shown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              --NGRWIPLLDWVGWNNWGYNGVVS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------GFTHADVT-AVPSTIFTSPVFTTG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  421 SPKLGWMDEAMAIDPFNSDRMLYGTGATLYA----
A;Status: nucleic acid sequence not shown; tra
A;Molecule type: DNA
A;Residues: 1-1749 <ARN>
A;Cross-references: EMBL:D90903; GB:AB001339;
A;Note: the nucleotide sequence was submitted
CGenetics:
A;Start codon: GTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     578 ASADG---SRFVWAPGDPGQPVVYAVGFGNSWA--
                                                                                                                                                                                                                                                                                                                                                                                              23
                                                                                                                                                                                                                                                                                                                                                                                                                                                   63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               786
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   499
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            φ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

ö

ä

695

951

```
811 SPSP---SPSSSPSSPSSPSP----SPSPSPSPSP----SRSP----SPSASPSPSS 848
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 --- SSSPSPTPSSSPVSGGVKVQYKNND 881
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               882 SAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNC--DWAAIGCGNIRA 939
SEPSPSSTQSSGAQSTLTTPSPNPSQSTSSLESSTSGATTSSGSAGTTMTSPSQSSSVGS 760
                                                                                                                                                                                              870 PYPSQSTSPVESSTTPS-----PGSPGTTLTSTSPSPSQSTTIGSTQGSTSPGIS 919
                                                                                                                                                                                                                                                                                                                                                               696 GVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQA 755
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  756 ITGDHANLRRVYIGTNGRGIVYGDIGGAPSGSPSPSVSPSASPSLSPSPS-----PSSSP 810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             540 FDPSSQPNDRHVAFSTDGGKNWFQGS--EPG-GVTTGGTVAASADGSRFVWAPG-----
                                                                                                                  761 SQGSTSP----AASTTSGEMTSQGSTQTPGSSVSTSAAILTSTQQSVSTNSPGSTVTRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  952 ------SSGSTVTVGSTEG--SSSPIPSTSQNTNPSTSSGSSMSTQTPQSSQ
                                                                                                                                                                  -----DPGQPVVY----AVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRS
                                                                                                                                                                                                                                                                   640 TDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSA---IT
                                                                                                                                                                                                                                                                                                                                                                                                                 920 TTSEEMTSQGSTQTPGSTGSTVTQPSTVSDST------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           09:13:57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2, 2002,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1160 STGSTVTQPSTGSGST 1175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               940 SFGS--VNPATPTADT 953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Search completed: July
Job time: 125 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 849 SPSPSSSP---
                      701
                                                                                                               g
                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                    Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                       ò
                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                hypothetical protein K06A9.1a - Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Accession: T34434
R; Geisel, C.; Gattung, S.
submitted to the EMBL Data Library, December 1996
A; Description: The sequence of C. elegans cosmid K06A9.
A; Reference number: 221525
A; Accession: T34434
A; Reference number: 221525
A; Residues: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Molecule type: DNA
A; Mosicule type: DNA
A; Map position: X
A; Introns: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 1088/1; 1367/1; 2039/1; 2075/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2075/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             449 VVSGSSSTYGSSTPSASSSAGTASTISGSTGSTATIVPGSS------SSVGSSTQSA 500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         339 TANDYFGYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWT----RIWDWTSYPNR 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SSTVPSSSSPQPSSQSPAPNTGSTTPSQTSSQSPSPSMNPSSSTPTGSSQSTITPEGSTA 646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SSPIGSTGSTESVATEVISQSI-----VPSGSSLGTQSINSSPSPSSLSPSTSGMSTLT 700
                      175 SPSPSPSPSPSPSPSPSPKASPSPSPKASPSPKASPASPAPSPQPSPTPSPKASP 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          235 SLGTSSSPLPSSISTSALPIASSSASSSPSAASSTTPVVLSSSTIQSSSGTFPSSVASSP 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               73 GAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAAD-----PINTNK 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             127 WAAAVGMYTNSWDPNDGAILR--SSDQGATWQITP-LPFKLGGNMP---GRGMGERLAVD 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              331 --STIGSSTPSASSSSGTMSTISGSTGSTVTVVPGSSSSTFASSTPIASSSSPGSTVTVA 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PNNDNILYFGAPSGKGL---WRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVV 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             279 VPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTD 338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      541 PAS-----TGSTITISGSSSIIVST-----VSGSTVSGSTGTSQSTLASSTATPGS 586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       396 SLRYVLDISAE------PWLTFGVQPNP---PVPSPKLGWMDEAMAIDPFNS 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DRMLYGTGATLYATNDLTKWDSGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDL- 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ------GGFTHADVTAVPS-----TIFTSPVFTTGTSVDYAELNPSIIVRAGS 539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    18 SLLAATASFAVAAALGVLPIA---ITASPAHAATTQPYTWSNVAI--GGGGFVDGIVFNE 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                295 STVG-----STSG-----AASSSSYATVSTIAGSTGSTITPVPGSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WVAFDK------SSSSLGQAS-----KTIFVGVADPNNPVFWSRDGGATWQA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      315;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 2232;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 5.7%; Score 290.5; DB 2; Best Local Similarity 21.8%; Pred. No. 1.6e-05; Matches 239; Conservative 114; Mismatches 428;
                                                                                                                     841 SASP--SPSSSPSSSSPSSSPSTPSSSP 868
                                                                                                                                                 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        498
                                                                    a
                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                      ò
                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ô
```

```
GUNV_ERWCA
RPB1_CREGE
RPB1_CARED
PCCA_CANFA
ALA1_CANAL
AFMU_PIG
ALS1_CANAL
RPB1_DICUI
RPB1_DICUI
RPB1_DICUI
RPB1_ARATH
CSP_PLAFA
505
467
1859
2333
1419
11150
1260
2364
412
1840
397
198.5
197.5
197.5
197.5
197.5
197.5
193.5
192.5
192.5
191.5
2, 2002, 09:11:52 ; Search time 28.05 Seconds (without alignments) 1321.020 Million cell updates/sec
                                                                             US-09-917-376-1
5135
1 MDRSENIRLTMRSRRLVSLL......RASFGSVNPATPTADTYLQX 957
       GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                    OM protein - protein search, using sw model
                                                   July
```

cricetulus carenchabdi cannor all manipular and alb sus scrofa candida alb candida alb dictyosteli bos taurus plasmodium arabidopsis

047096 (P11414 CP11414 CP11414 CP16356 CQ28343 CO1345 CP12021 SP46590 CP13608 PP13608 PP02893 PP19597 PP19597

ALIGNMENTS

RESULT 1 GGUNA_CALSA GGUNA_CALSA AC GGUNA_CALSA DT 01-AU DT 01-OC DT 15-UU DE Endog DD CCID OC Bacte O
--

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

105224

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_40:*

Database :

105224 seqs, 38719550 residues

BLOSUM62 Gapop 10.0 , Gapext 0.5

Scoring table:

Searched:

Title: Perfect score:

:uo

Sequence:

and is delived by amalysis of the total score distribution.	
COCAL	SUMMARIES
נו	SUMM
5	
ana t y a ra	
7	
מכודגפת	
9	

	criptio	P22534 caldocellum	P22533 caldocellum	34			-	-	P93329 medicago tr	-		Q06852 clostridium	P07897 rattus norv	P19274 thermoprote					P54583 acidothermu	P74745 synechocyst				-		P10475 bacillus su	P24928 homo sapien	PO8775 mus musculu	_	~	_	σ	364	Q9y4h2 homo sapien
SUMMARIES		GUNA_CALSA	MANB_CALSA	Y091_NPVOP	GUNB_CALSA	VTP3_TTV1V	GP1_CHLRE	CIPA_CLOIM	NO20_MEDTR	CIPB_CLOTM	GUNI_CLOTM	SLP1_CLOTM	PGCA_RAT	VTPX_TTV1	PGCA_MOUSE	TEGU_HSV11	APG_BRANA	SPG7_DICDI	GUN1_ACICE	SPKC_SYNY3	PGCA_HUMAN	SSGP_VOLCA	GP10_DICDI	APG_ARATH	GUNA_PAELA	GUN2_BACSU		RPB1_MOUSE	GUNN_ERWCA		GUNA_MICBI	GUN3_BACSU	AMYH_YEAST	IRS2_HUMAN
	ength DB	1742 1				474 1			268 1	772 1	σ.								562 1							499	970	70	44	66		499	1367 1	324
ø		7.7	9.9	5.7	5.6	5.5	5.3	5.1	4.9	4.8	4.8	4.8	4.6	4.3			4.2	4.2	4.1	4.1	4.1	4.1	4.0	4.0	4.0	4.0	4.0	4.0	4.0				3.9	3.9
	Score	397.5	340.5	293	290	283.5	272	261.5	250	245	245	244	235.5	222.5	221	218.5	216	214	210.5	209	209	208.5	207.5	205.5	205.5	204.5	204.5	204.5	204	203.5	201	199.5	199.5	199
	Result No.	1	7	m	4	S	9	7	ω,	0	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	53	30	31	32	33

```
538 TLSTNYNQGAKVSGPYVWDSSRNIYYILVDFTGTLIYPGGQDKYKKEVQFRIAAPQNVQ-
                          766 VYIGTNGRGIVYGDIGGAPSGS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; L01257; AAA71887.1; -.
                                                                                                                                                                                                                                           STANDARD:
                                                                                                                                                                                                                                                                                                                                                                                                                                               multidomain enzyme.
                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=44001;
                                                                                                                                                                                                                                                                01-AUG-1991
                                                                                                                                                                                                                                                                           01-JUL-1993
16-OCT-2001
                                                                                                                                                                                                                                           MANB_CALSA
                                                                                                                 868
                                                                                                                                                                                                                                  MANB_CALSA
                                                597
                                                                                                                                                                                                                                                                                                                                   MANA
                          ò
                                                QQ
                                                                                           g
                                                                                                                  ò
                                                                                                                                        a
                                                                                                                                                              ð
                                                                                                                                                                                   q
                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                   35;
                                                                                                                                                                                                                                                                                                                                                                     317 DVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQI-----SWW- 368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEETAVN 479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           621
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            501
                                                                                                                                                                                                                                                                                                                                                      262 PNNPVFWSRDGGATWQA----VPGAPTGFIPHKGVFDPVNHVLY-IATSNTGGPYDGSSG 316
                                                                                                                                                                                                                                                                                                                                                                                                                        EYKDAFVKSGQLEHI--LNQIEWVNDYF-----VKCHPSKYVYYYOVGDGGKDHAWWG 151
                                                                                                                                                                                                                                                                                                                                                                                                                                               -----PDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPPV 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   152 PAEVMOMERPSFKVTQSSPGSA-------VVAETAASLAAASIVLKDRN 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          225 GYYNSWSGFYDELSWAAVWLYLATNDSTYLTKAE----SYVQNWPKISGSNIIDYKWAH 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     280 CWDDVHNGAAL---LLAKITDKDTYKQIIESHLDYWTTGYNGERIKYTPKGLAWLDQWGS 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 669 KEGDLWLAASSGLYHSTNGGSSWSA-----ITGVSSAVNVGFGKSAPGSSYPAVF 718
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -- DQHQYGNWGQAITGDHANLRR 765
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              194 PTKAATYLQHA-----YTAAN 224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   337 LRYATTTAFLAFVYSDWSGCPTGKKETYRKFGESQIDYALGSTGRS--FVVGFGTNPPKR 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               395 PHHRTAHSSWADSQSIPSYHRHTLYGALVGGPGSDDSYTDDISNYVNNEVACDYNAGFVG 454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ------VNNQSGWPARATNKLKFRYFVDLSELIKAGYS+-----PNQL 537
                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----RVNP----KTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           455 ALAKMYLLYGGNPIPDFKAIETPTNDEFF-----VEAGINASGTNFIEIKAI--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DLISPPSG-----APLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----GKNW-----GKNW-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------FQGSEPGGVTTGGTVAASADG-SRFVWAPGDPGQPVVYAVGFG-----
                                                                                                                                                                      CELLULOSE-BINDING (BY SIMILARTYY).
LINKER ("HINGE") (PRO-THR BOX).
LELULULOSE-BINDING (BY SIMILARITY).
LINKER ("HINGE") (PRO-THR BOX).
CATALYTIC 2.
                                                                                                                                                                                                                                                                                                                                  Indels 303;
                                                                                                                 Signal; Repeat.
                                                                                                                                                                                                                                                                                                             Length 1742;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----NAQIRSD----
                                                                                                                                                             LINKER ("HINGE") (PRO-THR BOX).
                                                                                                                                                                                                                             BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
T -> A (IN REF. 2).
W; 3F0699A2123EED07 CRC64;
InterPro; IPR000556; Glyco_hydro_48.
InterPro; IPR001701; Glyco_hydro_9.
Pfam; PF00942; CBD_3;
Pfam; PF00759; Glyco_hydro_48; 1.
Pfam; PF00759; Glyco_hydro_9; 1.
PRINTS; PR00844; GLHYDRALSE48.
ProDom; PD011903; Glyco_hydro_48; 1.
ProDom; PD011903; Glyco_hydro_48; 1.
PROSTITE; PS00569; GLYCOSYL_HYDROL_F9_1; 1.
PROSTITE; PS005699; GLYCOSYL_HYDROL_F9_2; 1.
Cellulose degradation; Hydrolase; Glycosidase; Sic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --
                                                                                                                                                                                                                                                                                                             DB 1;
                                                                                                                                                                                                                                                                                                                                 Matches 189; Conservative 109; Mismatches 269;
                                                                                                                                                                                                                                                                                                          7.7%; Score 397.5; DB 1
21.7%; Pred. No. 4.9e-10;
                                                                                                                                        ENDOGLUCANASE A.
                                                                                                                                                   CATALYTIC 1
                                                                                                                             POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            719 VVGTI----GGVTGAYRSDDCGTTWVLIND-
                                                                                                                                                                                                                               BY
BY
T
                                                                                                                                                                                                                                                                            MΜ
                                                                                                                                                                                                                                                                           193696
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         _
_
__
                                                                                                                                      1742
642
                                                                                                                                                                                            1060
1112
1742
396
                                                                                                                                                                                                                                                                             ¥
                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                           643
701
858
904
1061
                                                                                                                                                                                                                              396
434
443
1545
                                                                                                                                                                                                                                                     ACT_SITE
CONFLICT
SEQUENCE
                                                                                                                                                                                                                              ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                         Query Match
Best Local $
                                                                                                                                      CHAIN
                                                                                                                             SIGNAL
                                                                                                                                                             DOMAIN
                                                                                                                                                                          DOMAIN
                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                         101
                                                                                                                                                                                                                                                                                                                                                                                                                                               369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          604
   g
                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CELSTUS.
-- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL HYDROLASES).
-- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY
596
                                                                                                                                                                                                                                  THE SECOND MOST ABUNDANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -!- CATALYTÍC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
linkages in cellulose.
-!- MISCELLANEOUS: THIS ENZYME IS MOST ACTIVE AT PH 6 AND 80 DEGREES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Beta-mannanase/endoglucanase A precursor [Includes: Mannan endo-1,4-beta-mannosidase A (EC 3.2.1.78) (Beta-mannanase) (Endo-1,4-mannanase); Endo-1,4-beta-glucanase (EC 3.2.1.4) (Cellulase)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
Bacteria; Firmicutes; Bacillus/Clostridium group;
Thermoanaerobacter group; Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POLYSACCHARIDES IN NATURE. CONTAINS TWO CATALYTIC DOMAINS WITH MANNANASE AND ENDOGLUCANASE ACTIVITIES.
CATALYTIC ACTIVITY: Random hydrolysis of 1,4-beta-D-mannosidic
                                                                                                                                                                                                                                                                                                                                 P-VSGGVKVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYN
                                                                 -- PSPSVSPSASPSLSPSPSPS
                                                                                                                              ---WDNSNDYSFQDIKGVSSGSVVKTKYIPLYDEDIKVWGEEPGTS-----GVSPTPTAS
                                                                                                                                                                                               part of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linkages in mannans, galactomannans, glucomannans, and galactoglucomannans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-93119139; PubMed-1476429;
Gibbs M.D., Saul D.J., Luthi E., Bergquist P.L.;
"The beta-mannanase from 'Caldocellum saccharolyticum'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (Rel. 16, Created)
(Rel. 26, Last sequence update)
(Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Appl. Environ. Microbiol. 58:3864-3867(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1331 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               J (FAMILY 44 OF GLYCOSYL HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                             927 CDWAAIGCGNIRASFGSVNPATPTADTYLQ 956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          761 -DWAQIGASNVTFKFVKLSSSVSGADYYLE 789
```

 \sim

```
| :: | || ||:
636 KFVKLSSSVSGADYYLE 652
                                                                                                                             01-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   48;
                                                                                                                                           01-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GUNB_CALSA
P10474;
                                                                                     Y091_NPVOP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local
                                                                                                010341;
                                                                    Y091_NPVOP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GUNB_CALSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
             q
                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                       29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SGAPLISALGDLGGFTHADVTAVPS -- TIFTSPVFTTGTSVDYAELNPSIIVRA -- - GSF 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DPSSQPNDRHVAFSTDG----GKNWFQGSEPGGVTTGGTVAASADGSRFVWAPGDPGQP 595
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              234 NTASKVEEYIKSFVDKGLPLVIGEFGHQHTD--------GDPDEE 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      426 WMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEETAVNDLISPP 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----AIKALRD-AGFKHTIMVDAPNWGQDWSNTMRDNAQSIMEADPLRNLVFSIHMYGVY 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       271 AIVRYAKQYKIGLFSWSWCGNSSYVGYLDMVNNWDPNNPTPW-----GQWYKTNAIGTSS 325
                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels 191; Gaps
                                                                                                                                                                                 BETA-MANNANASE/ENDOGLUCANASE A.,
CATALITIC (MANNANASE ACTIVITY).
PRO/SER/THR-RICH (PT BOX).
SUBSTRATE-BINDING (POTENTIAL).
PRO/SER/THR-RICH (PT BOX).
SUBSTRATE-BINDING (POTENTIAL).
PRO/SER/THR-RICH (PT BOX).
CATALITIC (ENDOGLUCANASE ACTIVITY).
PROTON DONOR (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  383 IRPWLKVVNSGSSSIDLSRVTIRYWYTVDGERAQSAISDWAQIGASNVTFKFVKLSSSVS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VV----YAVG-FGNSWAASQGVPANAQIRS--DRVNPKTFYALSNGTFYRSTDGG---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 326 TPTPTSTVTPTPTPTPTPTVTATPTPTPTPVSTPATSGQIKVLY---ANKETNSTTNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ---WL----AASSGL-----YHSTNGGSSWSAITGVS--SAVNVGF----GKSAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |: | | | : | | | : | | GADY--YLEIGFKSGAGQLQPGKDTGEIQMRFNKDDWSNYNQGNDWSWIQSMTS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----TPAPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            821 SSSPSPSPSPSPSPSPSASPSPSSSSPSSSSPSSSPSPSSSPVSGG-VKVQYKN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TATPTPTPTPTVTPTVTPTVTPTVTATP------TPTPTPTPTPVSTPATGGQIKVLYAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      880 NDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAIGCGNIRA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----VTFQPVAAGLP----SSGAVGVMFHAVPGKEGDL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        711 GSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDD---QHQYGN---WGQAITGDHANLR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RVYIGTNGRGIVYGD----IGGAPSGSPSPSVSPSASPSLSPSPSSSSPSPSPSPSSSSP
                                                                                                                                                                                                                                                                                                                                                                                                          DB 1; Length 1331;
                                                                                                                                                                                                                                                                                                                                                  5
                                                                                                                                                                                                                                                                                                                     NUCLEOPHILE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                TPTPTPT -> RQHQHRQ (IN REF. MW; FFBCA51BB8D8F0E0 CRC64;
                                                                                                           ProDom; PD001947; CBD_3; 2.
PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
Hydrolase; Glycosidase; Cellulose degradation; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                      97; Mismatches 189;
                                                                                                                                                                                                                                                                                                                                                                                                       6.6%; Score 340.5; DB 1
22.7%; Pred. No. 9.6e-08;
                                                                                                                                                                                                                                                                                                                                  T -> P (IN REF. 2)
TPTPTPT -> RQHQHRC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---YGENEKVTAYIDGVLVWGQEPSGA-----
                                    HSSP, Q06851, INBC.
InterPro: IPR001956; CBD_3.
InterPro: IPR001547; Glyco_hydro_F5.
Pfam: PF00942; CBD_3; 2.
Pfam; PF00150; cellulase; 1.
                                                                                                                                                                                                                                                                                                                                                               146892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 940 SFGSVNPATPTADTYLQ 956
                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                       Multifunctional enzyme
                                                                                                                                                                                                                                                                        780
            PIR; B43745; B43745.
PIR; A48954; A48954.
                                                                                                                                                                                                                                                                                                                                                               AA;
                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                 338
340
1331
                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches 140;
                                                                                                                                                                                                                                                                                                                                  CONFLICT
                                                                                                                                                                                                                                                                                                    ACT_SITE
                                                                                                                                                                                                                                                                                                                     ACT_SITE
                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                          Local
                                                                                                                                                                      SIGNAL
                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               495
```

ò 셤 ò g à q ò a

ò 셤 Q g δ g ò g ŏ 셤 ò g

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-JUL-1989 (Rel. 11, Created)
01-JUL-1989 (Rel. 11, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Endoglucanase/excglucanase/excglucanase
(EC 3.2.1.4) (Endo-1.4-beta glucanase) (Cellulase)
(Cellobiohydrolase); Excglucanase (EC 3.2.1.91) (Exocellobiohydrolase)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          78 SPITPSPILSPIPSPIPTPSPIPSPIPSPIPSPIPSPIPSPIPTPSPIPTPSPIPTPSPIPTP 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "The sequence of the Orgyia pseudotsugata multinucleocapsid nuclear polyhedrosis virus genome."; Virology 229:381-399(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
Bacteria; Firmicutes; Bacillus/Clostridium group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDINE-89098398; PubMed-2789517;
Saul D.J., Williams L.C., Love D.R., Chamley I.W., Bergquist P.I
"Nucleotide sequence of a gene from Caldocellum saccharolyticum
                                                                                                                                                                                 Orgyla pseudotsugata multicapsid polyhedrosis virus (OpMNPV).
Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-97271300; Pubmed-9126251;
Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
Rohrmann G.F.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 279;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    29289 MW; 6FA4DAA01009DBF0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5.7%; Score 293; DB 1;
55.8%; Pred. No. 1.9e-06;
tive 26; Mismatches 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -!- SIMILARITY: TO CORRESPONDING ORF IN ACMNPV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Thermoanaerobacter group; Caldicellulosiruptor.
                                                           (Rel. 35, Created)
(Rel. 35, Last sequence update)
(Rel. 35, Last annotation update)
1 29.3 kDa protein (ORF92).
279 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 1039 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       138 SPTPSPTPSPTPSPTPSPTPSP 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   843 SPSPSSSPSSSSPSSSPSSSP 868
      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (1,4-beta-cellobiohydrolase)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; U75930; AAC59091.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
      STANDARD;
                                                                                                                                                                                                                                                   Nucleopolyhedrovirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      l protein.
279 AA; 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                NCBI_TaxID=164623;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=44001;
                                                                                                                                                   Hypothetical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hypothetical
SEQUENCE 2
                                                              01-NOV-1997
```

ö

us-09-917-376-1.rsp

```
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                      Repeat
                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                       REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                    EPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GP1_CHLRE
  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7;
                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the FWBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                          linkages in cellulose.
-:- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glucosidic linkages in cellulose and cellotetraose, releasing cellobiose from the non-reducing ends of the chains.
-:- SIMILARITY: IN THE N-PERMINAL SECTION; BELONGS TO CELLULASE FAMILY F (FAMILY 10 OF GLYCOSYL HYDROLASES).
-:- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY IS SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WGQAITGDHANLRRVYIGTNGRGIVYGDIGGAPSGSPSPSSVSPSASPSLSPSPSSSPS 811
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             THE N-TERMINAL
                                      DOMAIN HAS EXOGLUCANASE ACTIVITY WHILE THE C-TERMINAL DOMAIN IS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PSPSPSSSPSSSPSPSPSPSPSPSRSPSPSASPSSSSPSPSSSSPSPSSPSPSSSPVSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  872 G-VKVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       THR/PRO-RICH, TANDEM REPEATS OF T-P. CELLCHOSE-BINDING (BY SIMILARITY). THR/PRO-RICH, TANDEM REPEATS OF T-P. PROTON DONOR (POTENTIAL).
                                                                CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5.6%; Score 290; DB 1; Length 1039; 33.0%; Pred. No. 1e-05; Live 43; Mismatches 59; Indels 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ENDOGLUCANASE/EXOGLUCANASE B.
                                                                                                                                                                                                                                                                                                              PIR; S02711; S02711.

R HSSP; OG6851; 1NBC.

InterPro: IPR0010956; CBD_3.

InterPro: IPR001000; Glyco_hydro_10.

InterPro: IPR001000; Glyco_hydro_F5.

Pfam; PF00142; CBD_3; 1.

Pfam; PF00131; Glyco_hydro_10; 1.

PRINTS; PR00134; GLHVDRLASE10.

PRODON; PR00134; GLYCOSYL_HVDROL_F10; 1.

R PROSITE; PS00599; GLYCOSYL_HVDROL_F5; 1.

R PROSITE; PS00659; GLYCOSYL_HVDROL_F5; 1.

R PROSITE; PS00659; GLYCOSYL_HVDROL_F5; 1.

R Cellulose degradation; Hydrolase; Glycosidase; Repeat;

M Whitifunctional enzyme; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUCLEOPHILE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    59; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY.
0E0378171594DDAE CRC64;
exocellulase and endocellulase activity.";
          Nucleic Acids Res. 17:439-439(1989).
-!- FUNCTION: THIS PROTEIN IS MADE UP OF TWO DOMAINS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                474 AA.
                                                                                                                                                                      A (FAMILY 5 OF GLYCOSYL HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     931 AIGCGNIRASFGSVNPATPTADTYLQ 956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        479 QIGASNVTFKFVKLSSSVSGADYYLE 504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                    EMBL; X13602; CAA31936.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     117641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1039
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       416
570
618
177
                                                  AN ENDOGLUCANASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
nes 68; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             29
376
417
571
177
285
792
1039
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 5
VTP3_TTV1V
ID VTP3_TTV1V
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACT_SITE
ACT_SITE
ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 812
ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | : :|: | |: :| |: | |: |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               811
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      581 DGSRFVW-----APGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGT 635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels 127; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 THR-PRO REPEATS REGIONS AND TWO NEAR IDENTICAL REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        48 DG---VWVIKLRSPLSPGQSITI-----TASSGTP-----NIDPTIALYYNNGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          636 FYRSTD-------GGVTFQPVAAG------LPSSGAVGVMFHAVPGKEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      672 DLWL-----AASSGLYHSTNGGSSWSAITG--VSSAVNVGFGKSAPGSSYPAVFVVGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               140 DRWATTPTSLDAVGLRLYADTN---DWFGVVRKYVNGAQNVSIEQKISG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             723 IGGVTGAYRSDDCGTTWVLIND-DQHQYGNWGQ-----AITGDHANLRRVYIGTNGRGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----PSVSPSASPSLSPSPSSSPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Neumann H., Zillig W.; "Nucleotide sequence of the variant "Nucleotide sequence of the viral protein TPX of the TTV1 variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           812 PSPSPSSSPSSPSPSPSPSPSPSRSPSPSASPSPSSSSPSSSSPSSSSPTPSSSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 474;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        89 SYSNLTLVGSPTVSIVQDFGGYAISAYASGDFFLVASPTGFTPSSSRLLVV---
                                                                                                                                                                                                            Thermoproteus tenax virus 1 (strain VT3) (TTV1).
Viruses; dsDNA viruses, no RNA stage; Lipothrixviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                THR-PRO(N).
1037109136598829 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5.5%; Score 283.5; DB 1
23.8%; Pred. No. 8.5e-06;
Live 61; Mismatches 84
                01-FEB-1991 (Rel. 16, Last sequence update) (Viral protein TPX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         THR-PRO(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            THR-PRO(N)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nucleic Acids Res. 18:2171-2171(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE=90245666; PubMed=2336394;
01-NOV-1990 (Rel. 16, Created)
01-NOV-1990 (Rel. 16, Last sequence)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GPI_CHLRE STANDARD;
09FPQ6; Q03927;
01-MAR-2002 (Rel. 41, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       49507 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; X14717; CAA32838.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               776 VYGDIGGAPSGSPS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PIR; S06686; S06686
PIR; S15921; S15921
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          447
474 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
Matches 85; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                     Lipothrixvirus.
NCBI_TaxID=10481;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NG I
```

```
CIPA_CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                .;
%
                                                                                                                                                                                                                                                                       PARTIAL PRELIMINARY SEQUENCE FROM N.A. MEDLINE-91017504; PubMed-1699225; Adair W.S., Apt K.E.; Adair W.S., Apt K.E.; "cell wall regeneration in Chlamydomonas: accumulation of mRNAs encoding cell wall hydroxyproline-rich glycoproteins."; Proc. Natl. Acad. Sci. U.S.A. 87:7355-7359(1990).
                                                                                                                                                                                                                         "Glycosylated polyproline II rods-with-kinks as a structural motif in plant hydroxyproline-rich glycoproteins."; Blochemistry 40:2978-2987(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PSGSPSPSVSPS-ASPSLSPSPSPSS---SPSPSPSPSS----SPSSSPSPS---PSPSPSPS 833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PSRSPSPSASPSSSPSSSPSSSPSSPTPSSSPVSGGVKVQYKNNDSAPGDNQIKPGL 893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PSPSPSPSPSPSPSPSPSPSPSPIPSPSPKPSPSPVA--VKLVWADDAIAFDD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VEGETATIVE CELL WALL PROTEIN GP1.
49 X 5 AA APPROXIMATE PPSPX REPEATS.
                                                                                               Eukarýota, Viridiplantae, Chlorophyta, Chlorophyceae, Volvocales,
Chlamydomonadaceae, Chlamydomonas.
                                                                                                                                                                                           Ferris P.J., Woessner J.P., Waffenschmidt S., Kilz S., Drees J.,
Goodenough U.W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLY-PRO.
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
01-MAR-2002 (Rel. 41, Last sequence update)
01-MAR-2002 (Rel. 41, Last annotation update)
Vegetative cell wall protein gpl precursor (Hydroxyproline-rich
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 555;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                41; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAIGC-GNIR 938
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6A584A90465502F5 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5.3%; Score 272; DB 1;
44.6%; Pred. No. 3.1e-05;
iive 15; Mismatches 41;
                                                                                                                                                                                                                                                                                                                                                                                                                       -!- SUBUNIT: Associates with GP2 and GP3.
-!- PTM: N-glycosylated and O-glycosylated.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR003882; Pistil_extensin.
InterPro; IPR002965; P. Ich_extensn.
PRINTS; PR01217; PRICHEXTENSN.
PRINTS; PR01218; PSTLEXTENSIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; AF309494; AAG45420.1; --
EMBL; M58496; AAA69706.1; ALT_SEQ.
GlycoSuiteDB; Q9FPQ6; --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Signal.
                                                                                  Chlamydomonas reinhardtii.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  54219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Glycoprotein; Repeat;
SIGNAL 1 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30
40
259
399
455
493
555 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                 NCBI_TaxID=3055;
                                                   (coprotein 1).
                                                                                                                                                                              PubMed-11258910;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Sim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
DOMAIN
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

RESULT

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 1820-1853 FROM N.A.
MEDLINE=93209931; PubMed=8458832;
Fujino T., Beguin P., Aubert J.-P.;
"Organization of a Clostridium thermocellum gene cluster encoding the cellulosomal scaffolding protein CipA and a protein possibly involved in attachment of the cellulosome to the cell surface.";
J. Bacteriol. 175:1891-1899(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN: CONTAINS 9 COPIES OF A DOMAIN (COHESIN) OF ABOUT 150 RESIDUES THAT BINDS TO THE DOCKERIN DOMAIN BORN BY THE CATALYTIC COMPONENTS OF THE CELLULOSOME.

DOMAIN: A 24 RESIDUES DOMAIN IS REPEATED TWICE IN THIS ENZYME AS WELL AS IN OTHER C.THERMOCELLUM CELLULOSOME ENZYMES. THIS DOMAIN MAY FUNCTION AS THE BINDING LIGAND FOR THE SL COMPONENT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-97238934; PubMed-9083107;
Shimon L.J.W., Bayer E.A., Morag E., Lamed R., Yaron S., Shoham Y.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-97076134; PubMed-8918451; TOFINO J., Lamed R., Chirino A.J., Morag E., Bayer E.A., Shoham Y.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               J. Mol. Biol. 273:701-713(1997).
-!- FUNCTION: ACTS AS A SCAFFOLDING PROTEIN IN THE CELLULOSOME. IT PROMOTES BINDING OF CELLULOSE TO THE CATALYTIC DOMAINS OF THE
                                                                                                                                                                                                                                                                                                                                                                               'Sequencing of a Clostridium thermocellum gene (cipA) encoding the
                                                                                                                                                                                 Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae; Clostridium.
                                     01-JUN-1994 (Rel. 29, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
01-MAY-2002 (Rel. 41, Last annotation update)
01-MAY-2002 (Rel. 41, Last annotation update)
cellulosomal scaffolding protein A precursor (Cellulosomal glycoprotein SI/SL) (Cellulose integrating protein A) (Cohesin).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "Crystal structure of a bacterial family-III cellulose-binding
                                                                                                                                                                                                                                                                                                                 MEDLINE-93302508; PubMed-8316083;
Gerngross U.T., Romaniec M.P.M., Kobayashi T., Huskisson N.S.,
                                                                                                                                                                                                                                                                                                                                                                                                     cellulosomal SL-protein reveals an unusual degree of internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "A cohesin domain from clostridium thermocellum: the crystal structure provides new insights into cellulosome assembly."; Structure 5:381-390(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 1213-1361.
MEDILINE-98022914; PubMed-9402065;
Tavares G.A., Beguin P., Alzari P.M.;
"The crystal structure of a type I cohesin domain at 1.7-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           domain: a general mechanism for attachment to cellulose.'
EMBO J. 15:5739-5751(1996).
                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A., AND SEQUENCE OF 29-40 AND 42-43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 184-321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 368-522.
  AA.
1853
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CELLULOLYTIC ENZYMES.
SUBCELLULAR LOCATION: CELL SURFACE.
                                                                                                                                                                                                                                                                                                                                                                                                                                               Mol. Microbiol. 8:325-334(1993).
                                                                                                                                                                   Clostridium thermocellum.
                                                                                                                                                                                                                                                                                             27405;
                                                                                                                                                                                                                             NCBI_TaxID=1515;
                                                                                                                                                                                                                                                                                                                                                             Demain A.L.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Steitz T.A.;
                                                                                                                                                                                                                                                                                               STRAIN=ATCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 resolution.
                                                                                                                                                                                                                                                                                                                                                                                                                             nomology.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Frolow F
```

us-09-917-376-1.rsp

```
933 -GCGNIRASFGSVNPATPTADTYLQ 956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         268 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-3880;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CIPB_CLOTM
Q01866;
Q1-OCT-1994 (
01-OCT-1994 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Simi
Matches 71;
                                                                                                                                          NO20_MEDTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                nodulins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                         ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6
                                                                                                                            NO20_MEDTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CIPB_CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    725
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                829
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                              ΩI
                                                     g
                                                                                                                                                               E A E E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   880
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------GVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG 706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KSAPGSSYP----AVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGD--H 760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                508 VPSTIFTSPVFTTG---TSVDYAELNPSIIVR----AGSFDPSSQPNDRHVA--FSTDGG 558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        559 KNWFQGSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQI 618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              134 KSAAAAPITL--LEVGAFADNDLVEISTTFVAGGVNLGSSVPTTQP---NVPSDGVVVEI 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         245 KSFDTAIYPDRKIIVFLFAEDSG-TGAY---------AITKDGVF 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                  CELLULOSOMAL SCAFFOLDING PROTEIN A. CELLULOSE-BINDING (BY SIMILARITY).
9 x 150 AA APPROXIMATE TANDEM REPEATS.
1-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      115 R--------GTYAITQDG------VFAT------IVATV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     189 GKVTGSVGTTVEIPVYFRGVPSK ---- GIANCDFVFRYDPNVLEIIGIDPGDIIVDPNPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           619 RSDRVNPKTFYALSNGTFYRS-----TDGGV-----TFQPVAAGLPSSGAV---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ANLRRVYIGTNGRGIVYGDIGGAPSGSPSPSVSPSASPSLSPSPSSSSPSPSPSSSSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              280 AKIRATVKSSAPGYITFDEVGGFAD-----NDLVEQKVSFIDGGVNVGNATPTKGATP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SSSPSPSPSPSPSPSPSPSPSASPSPSSSPSSSSPSPTPSSSPVSGCVKVQYKNN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAI-----
                                                                                   PDB; INBC; 26-SEP-97.
InterPro; IPR001956; CBD_3.
InterPro; IPR002102; Cohesin..
InterPro; IPR002102; Cohesin..
InterPro; IPR002108; EF-hand.
InterPro; IPR002048; EF-hand.
InterPro; IPR003048; EF-hand.
InterPro; IPR00404; Dockerin_1; 2.
Prodom; PD001947; CBD_3; 1.
PROSITE; PS000408; EF_HAND; UNKNOWN_1.
PROSITE; PS000408; EF_HAND; UNKNOWN_1.
PROSITE; PS000408; CELLULOSOME_RPT; 2.
Cellulose degradation; Cell wall; Glycoprotein; Repeat; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24 AA APPROXIMATE REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       . AA (IN REF. 1).
3ABDDC03ABFC5372
 NOT ANNOTATED CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2 - 1 - 8 .
2 - 1 - 9 .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ļ
                                                                                                                                                                                                                                                                                                                                                                                                          AA; 196831
                  CAA47840.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1526
1690
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1847
                                                                                                                                                                                                                                                                                                                                                                                                                                                            1031
1196
                                                  PDB; 1ANU; 23-JUL-97
PDB; 1AOH; 08-JUL-98
                EMBL; X67506; CAA478
PIR; S36859; S36859.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity
nes 117; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1791
1791
1824
1615
1853
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             054
                                                                                                                                                                                                                                                                                                      3D-structure
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                   SIGNAL
CHAIN
DOMAIN
DOMAIN
REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
              EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  099
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               761
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         881
DORWELL STATE STAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ilensib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Spermatophyta; Magnoliophyta, eudicotyledons; core eudicots, Rosidae; eurosids I; Fabales, Fabaceae; Papilionoideae; Trifolieae; Medicago.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              780 IGGAPSGSPSPSVSPSASP-----SLSPSPSPSSSPSPSPSPSSSPSSSPSPSP---- 828
132 VLSSPPPPPSPPTPRSSTPIPHPPRRSLPSPPSPSPSPSPSPSPSPSPSPSPSPSPIPHPRRRS 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         early
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDH----ANLRRVYIGTNGRGIV-YGD 779
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                84 GIRGEHVDHYDGNTMVVLKK------TGIHHFISGKKRHCRLGLKLAVVVMVAP 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        192 PASPSPSPSLSKSPSPSESPSLAPSPSDSVASLAPSSSPSDESPSPAPSPSSSGSK 247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
STRAIN-CV. JERALONG;
MPDLINE-99187258; Pubmed-9526510;
Greene E.A., Erard M., Dedieu A., Barker D.G.;
"MtENOD16 and 20 are members of a family of phytocyanin-related of the contraction of the contraction
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 --SPSPSPSRSPSASPSPSSSPSPSSS----PSSSP---SPTPSSSPVSGGVK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 268;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6FA29C5798C75F91 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Plant Mol. Biol. 36:775-783(1998).
-!- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4.9%; Score 250; DB 1;
40.3%; Pred. No. 0.00012;
tive 14; Mismatches 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BY SIMILARITY.
EARLY NODULIN 20.
PLASTOCYANIN-LIKE.
                                                                                                                                                                                                                                                                                                                                                                                                                                               15-DEC-1998 (Rel. 37, Created)
15-DEC-1998 (Rel. 37, Last sequence update)
15-DEC-2099 (Rel. 38, Last annotation update)
Barly nodulin 20 precursor (N-20)
                                                                                                                                                                                                                                                                                                                                              268 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (Rel. 30, Created)
(Rel. 30, Last sequence update)
                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR003245; Cu_bind_like. Pfam; PF02298; Cu_bind_like; 1. ProDom; PD003122; Cu_bind_like; 1. Nodulation; Signal.
436 GITSNVKGTFVKMSSSTNNADTYLE 460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Medicago truncatula (Barrel medic)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28668 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; X99467; CAA67830.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22
268
?
```

```
g
                                                                                                               ŏ
                                                                                                                                         g
                                                                                                                                                                                              a
                                                                                                                                                                                                                                                   g
                                                       ò
                                                                                                                                                                    ð
                                                                                                                                                                                                                           δλ
                                                                                                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            14;
                                                                                                                                                                                                                                                                                                                                                                                               modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                       "Identification of the cellulose-binding domain of the cellulosome subunit S1 from Clostridium thermocellum YS.";

FEMS Microbiol. Lett. 78:181-186(1992).

-!- FUNCTION: ACTS AS ASCAFFOLDING PROTEIN IN THE CELLULOSOME. IT PROMOTES BINDING OF CELLULOSE TO THE CAPALYTIC DOMAINS OF THE CELLULOING OF THE NINE REPEATED DOMAINS WITH THE DUPLICATED SEGNENT OF 24 AMINO ACIDS PRESENT IN CATALYTIC SUBUNITS OF THE CELLULOSOME.
-!- SUBCELLULAR LOCATION: CELL SURFACE.
-!- SUBCELLULAR LOCATION: CELL SURFACE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   634 GTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPG-KEGDLMLAASSGLYHSTNGGSSWS 692
16-OCT-2001 (Rel. 40, Last annotation update)
Cellulosomal scaffolding protein B (Cellulosomal glycoprotein S1/SL)
(Cellulose integrating protein B) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ||: : || : | : | : | : || GTVAITQDG--VFAITVAIVKSAAAAPITLLEVGAFRDNDL------VEIST-----
                                                    Clostridium thermocellum.
Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
Clostridium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. 0.0006;
3; Mismatches 116; Indels 110;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 X 147 AA APPROXIMATE REPEATS.
1 (PARTIAL).
2.
                                                                                                                                            MEDLINE-93146373; PubMed-1490597;
Poole D.M., Morag E., Lamed R., Bayer E.A., Hazlewood G.P.,
Gilbert H.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4.8%; Score 245; DB 1; Length 772; 24.4%; Pred. No. 0.0006;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 24 AA APPROXIMATE REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BBF06DE5E094FE10 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRO/THR-RICH.
CELLULOSE-BINDING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRO/THR-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRO/THR-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro: IPR001956; CBD_3.
InterPro: IPR002102; Cohesin.
InterPro: IPR002105; Dockerin_1.
InterPro: IPR002048; EF-hand.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              766 2-733 2-766 2-82491 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00942; CBD_3; 1.
Pfam; PF00963; Cohesin; 3.
Pfam; PF00404; Dockerin_1; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; X68233; CAA48312.1; -. HSSP; Q06851; 1NBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 <1
    462
    462
    81
241
273
710
710
710
724
772 AA;</pre>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                      SEQUENCE FROM N.A.
                                                                                              NCBI_TaxID=1515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Simi
Matches 95;
                                                                                                                                      STRAIN-YS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28
 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
```

693 AITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNW 752

--TFVAGGVNLGSSVPTTQPNVPSDGVVVEIGKVTGS----VGTT---VEIPVYFRGVP 121

72

oy D

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
01-NOV-1995 (Rel. 32, Last annotation update)
Endoglucanase I precursor (EC 3.2.1.4) (EGI) (Endo-1,4-beta-glucanase)
                                                                                                                                                                                                                                            932
                                                                                                                                                                                      GGVNVGNATPTKGATPTNTATPTKSA----TATPPGHSVPTNTPTNTPANTPVSGNLKVE 284
                                                                                                                                                                                                                                                                                FYNSNPSDTTNSINPOFKVTNTGSSAIDLSKLTLRYYYTVDGQKDQTFW-CDHAAIIGSN 343
                                     SKGI----ANCDEVFRYDPNVLEIIGIDPRSIIV-----DPNPTKSFDTAIYADRK 168
                                                                                                                 IIVFLFCGRQRNRSVSITKDGVFAKIRATVKSSAPAYITFDEVGGFADNDLVEQKVSFID 228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "Gene sequence and properties of Cell, a family E endoglucanase from Clostridium thermocellum.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             J. Gen. Microbiol. 139:307'316(1993).

-! FUNCTION: THIS ENZYME CATALYZES THE ENDOHYDROLYSIS OF 1,4-BETA-GLOCOSIDIC LINKAGES IN CELLULOSS. LICHENIN AND CEREAL BETA-D-GLUCONIS. PRINCIPALLY ACTIVE AGAINSY BARLEY BETA-GLUCCAN.
-! CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic linkages in cellulose.
-! PATHWAY: CELLULOSE DERRADATION.
-! PATHWAY: CELLULOSE DERRADATION.
-! SIMILARIII: BELONGS TO CELLULASE FAMILY E (FAMILY 9 OF GLYCOSYL
                                                                                                                                                                                                                                        877 YKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAI----
                                                                                                                                                             Clostridium thermocellum.
Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
Clostridium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-93171873; PubMed-8436949;
Hazlewood G.P., Davidson K., Laurie J.I., Huskisson N.S.,
  -----IGTNGRGIVYGDIGGAPSGSPSVS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR001956; CBD_3.
InterPro; IPR001956; CBD_3.
InterPro; IPR001701; Glyco_hydro_9.
Pfam; PF00049; CBD_3; 2.
Prodom; PF0001947; CBD_3; 1.
PROSITE; PS00592; GLYCOSYL_HYDROL_F9_1; 1.
PROSITE; PS00698; GLYCOSYL_HYDROL_F9_2; 1.
Callulose degradation; Hydrolase; Glycosidase; Signal.
SIGNAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    879 AA
                                                                               -----PSASPSLSPSPSSSSPSPSPSSSPS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A., AND SEQUENCE OF 56-69
                                                                                                                                                                                                                                                                                                                                               1:::| ||||||:
344 GSYNGITSNVKGTFVKMSSSTNNADTYLE 372
                                                                                                                                                                                                                                                                                                                          ----GCGNIRASFGSVNPATPTADTYLQ 956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; L04735; AAA20892.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
  GQAITGDHANLRRVY - -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PIR; A47704; A47704.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10682;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HSSP; P26221; 1TF4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=1515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (Cellulase I).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gilbert H.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-NCIB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GUNI_CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 10
GUNI_CLOTM
753
                                       122
                                                                             794
                                                                                                                     169
                                                                                                                                                             822
                                                                                                                                                                                                                                                                                                                        933
```

us-09-917-376-1.rsp

```
t. It is produced through a collaboration icinformatics and the EMBL outstation-tute. There are no restrictions on its as long as its content is in no way tremoved. Usage by and for commercial ment (See http://www.isb-sib.ch/announce/ib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          44;
ce update)
.ion update)
sor (Outer layer protein B) (S-layer
                                                                                                                                                                                                                                           imocellum gene cluster encoding the LipA and a protein possibly involved to the cell surface.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  244; DB 1; Length 1664;
No. 0.0015;
smatches 316; Indels 292; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NI--LYFGAPSGKGLWRSTDSGATWSQMTNF 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NKVWAAVGMYTNSWDPNDGAILRSSDQGATW 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |:| : : | |: : 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KSSSSLGQASKTIFVGVADPNNPV----FWS 269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FDPVNHVLYIATSNTGGPYDGSSGDVWKFSV 323
                                                                                                              stridium group; Clostridiaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITAL.
SURFACE GLYCOPROTEIN 1.
156 AA APPROXIMATE REPEATS.
                                                                                                                                                                                                                                                                                                                                  AYERED CRYSTALLINE ARRAYS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              XIMATE TANDEM REPEATS OF

1-D-E-P.

RO/SER/THR-RICH.

(INCOMPLETE).
                                                                                                                                                                                                                                                                                                                                                                   HOMOLOGY (SLH) DOMAINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (INCOMPLETE).
F396695BA9FE74B CRC64;
```

```
01-FEB-1991 (Rel. 17, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
Aggrecan core protein precursor (Cartilage-specific proteoglycan core protein) (CSPCP).
                                                                                                                                                                                                                                                                                                                       740
                                                                                                                                   540
                                                                                                                                                                                               622
                                                                                                                                                                                                                   632
                                                                                                                                                                                                                                       959
                                                                                                                                                                                                                                                           678
                                                                                                                                                                                                                                                                                                    724
                                                                                                                                                                                                                                                                                                                                            784
                                                                                                                                                                                                                                                                                                                                                               764
                                                                                                                                                                                                                                                                                                                                                                                   ----SGSPSPSVSPSASPSLSPSPSSSPSPS--PSPSSSPSSSPSP-SPSPSPSRS 837
                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-88087070; PubMed-3693370; Doege K., Sasaki M., Horigan E., Hassell J.R., Yamada Y.; Complete primary structure of the rat cartilage proteoglycan core protein deduced from CDRA clones."; J. Biol. Chem. 262:17757-17767(1987).
                             371 EPGEEPTEEPVP-TETPVD-------PTPTVTEEPVPSELPDSYVIMELD--KT
                                                              437 NSDRMLYGT--GATLYATNDLTKWDSGGQIHIAPMVKGL---EETAVNDLISPPSGAPLI
                                                                                                              TGGTVLQSDNYGKTTAVANDV----GAGIINFAEAYSNLTKYRETGV-----AEET
                                                                                                                                 492 SALGDLG------GFTHADVTAVPSTIFTSPVFT-TGTSV-DYAELNPSIIVRAGSF
                                                                                                                                               DPSSQP-------NDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAA
                                                                                                                                                                                            573 EPTEEPVPTETPVDPTPTVTEEPVPSELPDSYVIMELDKTK------VKEGDVIIA
                                                                                                                                                                                                                                      623 TIRVNNIKNLAG------YQIGI-------KYDPKVLEAFNIETGDP
                                                                                                                                                                                                                                                          633 ------NGTFYRSTD---GGVTFQPVAAGLPSSGAVGVMF--HAVPGKEGDLWLAAS
                                                                                                                                                                                                                                                                       657 IDEGTWPAVGGTILKNRDYLPTGVAINNVSKGILNFAAYYVYFDDYREEGKSED-----
                                                                                                                                                                                                                                                                                                  679 SGLYHSTNGGSSWSAITGVSSAVNVGFG------KSAPGSSYPAVFVVGTIG
                                                                                                                                                                                                                                                                                                                       ------TGIIG--NIGFRVLKAEDTTIRFEELESMPGS------
                                                                                                                                                                                                                                                                                                                                          725 GVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDIGGAP
                                                                                                                                                                                                                                                                                                                                                              -IDGTYMLD----YVVIQPAP
        TSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQI-SWWPDTIIFRSTDGGAT
                                                 383 WTRIWDWISYP-----NRSLRYVLDISAEPWLIFGVQPNPPVPSPKLGWMDEAMAIDPF
                                                                                                                                                                                                                  SADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REVISION TO 698..
Doege K., Sasaki M., Horigan E., Hassell J.R., Yamada Y.;
J. Biol. Chem. 263:10040-10040(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                           838 PSPSASPSSSSPSSSSPSSSPSTPSSSPVSGGVKVQYKNNDSAPGD
                                                                                                                                                                                                                                                                                                                                                                                                                                      2124 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 1856-2124 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Rattus norvegicus (Rat).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AGC1 OR AGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PGCA_RAT
P07897;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12
                                                                                                                                                                          541
                                                                                                                                                                                                                  579
                                                                                                                                                                                                                                                                                                                                                                                   785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PGCA_RAT
                                                                                                                                                    a
                                                                                                                                                                                           g
                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   d
                                                                                          ŏ
                                                                                                                                  ò
                                                                                                                                                                          ò
                                                                                                                                                                                                                 Óγ
                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                          ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                         SIMILARITY).

DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO
DOMAIN: TWO GLOBULAR RECION, G3,
THE RECONTERNINGS G1 CONTAINS LINK DOMAINS AND THUS
CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS
THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS)
AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                THE CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE CHAINS, N-LINKED AND O-LINKED OLIGOSACCHARIDES.
SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
SIMILARITY: CONTAINS 4 LINK DOMAIN.
SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
                SWART; SM00032; CCP; 1.
SWART; SM00034; CLECT; 1.
SWART; SM00034; CLECT; 1.
SWART; SM00046; IGV; 1.
PROSITE; PS00290; IG_MHC; 1.
PROSITE; PS001241; LINK; 4.
PROSITE; PS00015; C_TYPE_LECTIN_1; 1.
PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
Glycoprotein; Cartilage; Proteoglycan; Lectin; Signal; Sushi; Repeat; Immunoglobulin domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AGGRECAN CORE PROTEIN. IG-LIKE V-TYPE DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LECTIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Interpro; IPR003006; Ig_MHC.
Interpro; IPR003596; Ig_v.
Interpro; IPR000338; Link.
Interpro; IPR00334; SGXXSG.
Interpro; IPR001304; Sushi_SCR_CCP.
Interpro; IPR001304; lectin_c.
Pfam; PP0059; lectin_c; I-
Pfam; PP02339; SGXXSG; S8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LINK 2.
LINK 4.
LINK 4.
C-TYPE I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LINK 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                G1-A.
G1-B.
G1-B'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; M13518; AAA41836.1; -.
EMBL; J03485; AAA21000.1; ALT_SEQ
MEDLINE=86250698; PubMed=2424893;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ProDom; PD000918; Link; 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00084; sushi; 1.
Pfam; PF00193; Xlink; 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    683
2036
2098
140
247
349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PIR; A23835; A23835.
PIR; A28452; A28452.
HSSP; P98066; 1TSG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             AND G3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
```

587 WAPGDPGQPVVYAVGEGNSWAASQGVPAN-AQ 1526	Qy 928 DWAAIGCGNIRASFGSVNPATPTADT 953 1 :	RESULT 13 VIPALITY1 ID VITX_ITY1 ID VITX_
FT DOMAIN 486 580 G2-B. FT DOMAIN 682 G2-B. FT DOMAIN 801 1226 CS-1. FT DOMAIN 801 1226 CS-1. FT DOMAIN 801 1226 CS-1. FT DOMAIN 1227 1309 CS-2. FT DOMAIN 1227 1309 CS-2. FT DISULEID 27 346 BY SIMILARITY. FT DISULEID 297 318 BY SIMILARITY. FT DISULEID 297 318 BY SIMILARITY. FT DISULEID 539 580 BY SIMILARITY. FT DISULEID 531 652 BY SIMILARITY. FT DISULEID 594 BY SIMILARITY. FT FT DISULEID 204 BY SIMILARITY. FT FT DISULEID 204 BY SIMILARITY. FT FT DISULEID	Query Match 4.6%; Score 235.5; DB 1; Length 2124; Best Local Similarity 21.3%; Pred. No. 0.0044; Matches 224; Conservative 115; Mismatches 359; Indels 353; Ga	QY 112 NGVVSTAADPINTNKWWAAVGMYTNSWDPNDCALLKSSDQGATWQITPLPFKLGGNMPGR 171 Db 979 SGIEDIGVLPGSEPTSAGGVDLGGLESGGESLETSAGGVE-DYDQFTERGGLETSA 1037 Cy 172 GMGERLAVDPNNDNIL

```
EMBL; L07049; AAC37670.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00084; sushi; 1.
Pfam; PF00193; Xilnk; 4.
Probom; PD000918; Link; 4.
SMART; SM00032; CCP; 1.
SMART; SM00034; CLECT; 1.
SMART; SM00406; IGV; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF00059; lectin_c; 1
Pfam; PF02339; SGXXSG; 62.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              247
349
381
581
683
2044
2106
140
247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2132
                                                                                                                                                                                                                                                                                                                                                                                               P98066; 1TSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                HSSP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
            +
                                                                                                                   -
17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ol-NOV-1997 (Rel. 35, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
Aggrecan core protein precursor (Cartilage-specific proteoglycan core
protein) (CSPCP).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   the aggrecan gene.";
Nat. Genet. 7:154-157(1994).
Nat. Genet. 7:154-157(1994).

-!- FUNCTION: THIS PROTBOGLYCAN IS A MAJOR COMPONENT OF EXTRACELLULAR MATRIX OF CARTILAGENOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEIN IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO HYALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR EGGION. MAY PLAY A REGULARORY ROLE IN THE MATRIX ASSEMBLY OF THE CARTILAGE.
-!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX (BY
                                                                                                                                                                                                                                                              815
                                                                                                                                                                                                                                                                                                                         290 PTXDITYVVFDVTPSPTPTPTPTTPTPTPTPTYDITYVIFDVTPSPTPTPTPTPTPT 349
                                              Gaps
                                                                   DGSRFVW-----APGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGT 635
                                                                                                                  636 FYRSTD-------GGVTFQPVAAG-----LPSSGAVGVMFHAVPGKEG 671
                                                                                                                                                                                                              775
                                                                                                                                                                                                                                     -----TYSVVNEIDISOFAAFTDPLVMYLSINGSTANV-KVYKQGSNIGT 229
                                                                                                                                                                                                                                                                                                           "Mouse cartilage matrix deficiency (cmd) caused by a 7 bp deletion in
                                                                                            88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-BALB/C; TISSUE-Cartilage;
MEDLINE-93104847; Pubmed-780622;
Walcz E., Deak F., Erhardt P., Coulter S.N., Fueloep C., Horvath P.
Doege K.J., Glant T.T.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Watanabe H., Kimata K., Line S., Strong D., Gao L.-Y., Kozak C.A.,
                                                                                                                                                                                                                                                                            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
NCBL_TaxID=10090;
                                                                                          48 DG---VWVVKLRSPLSPGQSITI-----TASSGTP-----NIDPTIALYYNNGS
                                                                                                                                                                672 DLWL-----AASSGLYHSTNGGSSWSAITG--VSSAVNVGFGKSAPGSSYPAVFVVGT
                                                                                                                                                                                     140 DRWATTPTSLDAVGLRLYADTN -- - DWFGVVRKYINGAQNVSIEQKISG -----
                                                                                                                                                                                                              723 IGGVTGAYRSDDCGTTWVLIND-DQHQYGNWGQ-----AITGDHANLRRVYIGTNGRGI
                                                                                                                                                                                                                                                             primary structure, chromosomal
                                              Indels 137;
                       Length 360;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Complete coding sequence, deduced primary structure, chrome localization, and structural analysis of murine aggrecan."; Genomics 22:364-371(1994).
                                                                                                                                        89 SYSNLTLVGSPTVSIVQDFGGYAISAYASGDFFLVASPTGFTPSSSRLLVV-
                     Query Match
4.3%; Score 222.5; DB 1;
Best Local Similarity 22.5%; Pred. No. 0.0025;
Matches 83; Conservative 55; Mismatches 94;
                                                                                                                                                                                                                                                                                                                                                                                                                                            2132 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN=129/SV;
MEDLINE=95004579; PubMed=7920633;
                                                                                                                                                                                                                                                                                                                                                                                                                                          PGCA_MOUSE STANDARD; F 061282; 064021; 01-NOV-1997 (Rel. 35, Last sequ 01-NOV-1997 (Rel. 35, Last sequ 01-NOV-1997 (Rel. 35, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 211-326 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mus musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                         862 PTPSSSPVS 870
                                                                                                                                                                                                                                                                                                                                                                                350 PTPTSTTSS 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Yamada Y
                                                                    581
                                                                                                                                                                                                                                                                                                                                                                                                                                PGCA_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                                       셤
                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                ò
                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                           ð
```

```
between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration
SIMILARITY).

DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO TERMINUS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3, MAKES UP THE COOH TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS) AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2 AND G3.
                                                                                                                                                                                                                                                                             THE CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE CHAINENS, N-LINKED AND O-LIORGASCHARTEDES.

DISEASE: DEPECTS IN AGC1 ARE THE CAUSE OF CARTILAGE MATRIX
DEFICIENCY (CMD), AN AUTOSOMAL RECESSIVE SYNDROME CHARACTERIZED BY CLEFT PALAFE, SHORT LIMBS, TAIL AND SNOUT: MUTATION IN STRAIN CMD CAUSES ABSENCE OF AGGRECAN BY TRUNCATION OF THE PROTEIN (MUTATION IN THE G1 DOMAIN).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -i- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
-i- SIMILARITY: CONTAINS 4 LINK DOMAINS.
-i- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
-i- SIMILARITY: CONTAINS 1 C-YPE LECTIN FAMILY DOMAIN.
-i- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
-i- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Glycoprotein; Cartilage; Proteoglycan; Lectin; Signal; Sushi; Repeat; Immunoglobulin domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AGGRECAN CORE PROTEIN. IG-LIKE V-TYPE DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C-TYPE LECTIN.
SUSHI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PROSITE; PS00290; IG_MHC; 1.
PROSITE; PS01241; LINK; 4.
PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MGD; MGI:99602, Agc.
Interpro; IPR003006: Ig_MHC.
Interpro; IPR003596; Ig_V.
Interpro; IPR003538; Link.
Interpro; IPR003324; SGXXSG.
Interpro; IPR003344; SGXXSG.
Interpro; IPR001304; lectin_C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LINK 1.
LINK 2.
LINK 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LINK 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            G1-A.
G1-B.
G1-B'
G2-B.
G2-B'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CS-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; S73722; AAB32160.1; -. EMBL; S73721; AAB32160.1; JOINED.
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                 ----SSTIDVSGESIAATIGSGES---SG 1668
                                                                                                                                                           LPEVILNISELVEGVIEPIVSQELG------HGPSMT-----YISRL 1704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEGLINE-8827437; PubMed-2839594;
MCGeoch D.J. Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
MCGeoch D.J. Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
MCGeoch D.J., Scott J.E., Taylor P.;
The complete DNA sequence of the long unique region in the genome of
herpes simplex virus type 1.";
J. Gen. Virol. 69:1531-1574(1988).
-!- FUNCTION: TEGUMENT PROTEIN.
-!- FUNCTION: TEGUMENT PROTEIN.
-!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL36,
EHV-1 24, EBV BPLF1, HVS-1 64, VZV 22, AND HCMV UL48.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2074 ASMDAAAELYVALVATTLTREFGC-----RWAQLEWGGDAAAPGPPLGPQSSTRHRVSFN 2128
                                                                                                                                                                                                                                                                                                                                      1765 DLHGITSAFHETDLEMTTPSTEVNSNPWTFQEGTREGSAAPEVSGESSTTSDIDTGTSGV 1824
                                                                                                              714 YPAVFVVGT--IGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIG-- 769
                       PSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSS 713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNWGY---NGVVSIAADPINTNKVW 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2129 ENDVLVALVASSPEHIYTF.----W--RLDLVRQHEYMHLTLPRAFQNAADSM-----
                                                                                                                                                                                                       --TNGRGIVYGDIGGAPSGSPSPSVS---PSASPSLSPSPSSSPSPSPSSSPSSSP
                                                                                                                                                                                                                                               1705 SEASGDASASGDLGGAVTNFPGSGVEASVPEASSDLSAYPEAGVGVSAAPEASSKLSEFP
                                                                                                                                                                                                                                                                                            427;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 3164;
                                                                                                                                                                                                                                                                                                                                                                                                                             1825 PSATPMASGDRIEISGEWS---DHTSE----VNVAIS----STITESEW 1862
                                                                                                                                                                                                                                                                                                                                                                                     PSSSPVSGGVKVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYW 913
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             35 X 2 AA TANDEM REPEATS OF IW; CC5D31FF4F9FE3F4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24 ASFAVAAALGVLPIAITASPAHAATTQPYTWSNVAIGGGGFVDG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Herpes simplex virus (type 1 / strain 17).
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.
NCBL_TAXID=10299;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative 100; Mismatches 402;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-MAR-1989 (Rel. 10, Created)
01-MAR-1989 (Rel. 10, Last sequence update)
01-APR-1993 (Rel. 25, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. 0.035;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Large tegument protein (Virion protein UL36)
UL36.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3164 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4.3%; Score 218.5; 20.3%; Pred. No. 0.0
                                                                    ------PYFSGDF-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2911 2980 35
3164 AA; 335857 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; X14112; CAA32311.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PIR; I30085; WMBEH6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
Matches 236; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TEGU_HSV11
                                                                    PSS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   P10220
                                                                 1637
                                                                                                                                                                                                     770
                                                                                                                                                                                                                                                                                                                                                                                     864
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TEGU_HSV11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     72
                       654
                                                                                                                                                                                                                                                                                            825
                                                               g
                                                                                                              δ
                                                                                                                                                        qq
                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1105 GYVSGIPSGGDGTET----SASGVEDVSGLPSGGEGLETSASGVEDLGPSTRDSLETSAS 1160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1307 GSGEV-----SGL----PLGELDSSAD---ISGL-----PSGTELSGQASGSPDS-- 1344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1345 SGETSG-------EETSGI 1371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1429 GISDLSGQPSGFPVFSGTATRTPDLASGTISGSGESSGITFVDTSFVEVTFTFREEEGL 1488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1591 EAGEGPSGILEFSGAHSGTP---DISGELSGSLDLSTLQSGQMETSTE------T 1636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1216 VGSASGALDFGKLPPGTLGSGQTPEVNGFPSGFSG-----EYSGADIGSGPSS---GLP 1266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGA-----TWQITPL 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         161 PFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPD----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1161 GVDVTGFPSGRGDPETSVSGVGDD---FSGLPSGKEGLETSASGA--EDLSGLPSGKEDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ------VGTYIANPT-DTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DPFNSDRMLYGTGATLYATNDLTKWD--SGGQIHIAPMVKGL--------EET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---SPVFTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AVGFGNSW-----AASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVAAGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             477 AVNDLISPPSGAPL------ISALGDLGGFTHADVTAVPSTIFT-----
                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    tch 4.3%; Score 221; DB 1; Length 2132; al Similarity 20.9%; Pred. No. 0.018; 199; Conservative 102; Mismatches 311; Indels 338;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    OB2BCDFC6CBDA163 CRC64;
              BY SIMILARITY.
CHOUSE GLONG.
N-LINKED GLCNAC.
N-LINKED GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CELL ATTACHMENT SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1489 GSVELSGFPSGETELSGTSGTVDVSEQSSGAIDSSGLTSPTPE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QGSEPG-----GVTTGGTVAASA-DGSRFVWAPGDP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    222008
                                                                                                                                                                                                                          1933
2042
2034
2092
2105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AA;
                                                                                                                                                                                                                                                                                                                                                         239
333
387
611
667
1675
                                                                                                         DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                     DISULFID
                                            DISULFID
                                                                 DISULFID
                                                                                    DISULFID
                                                                                                                                                                                                                                                                                          DISULFID
                                                                                                                                                                                                                                                                                                               DISULFID
                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              563
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

58;

Gaps

à	2y 129 AAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLA	1	78
g	::	7	2219
č q	2y 179 -VDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDV		216 2276
λά	217		274
QQ	2277		2306
δy	275		298
qq	Db 2307 SPPAPPNALYRPTGQHVAVPAAATHRTPAARVTAMDLVLAAVLLGAPVVVALRNTTAF		2364
ž ć	299VLYIATSNTGGPYDGSSGDVWKFSVTSGTW		340
3 8	341 NDYFGYSGLTIDROHPNTIWVATOISWWDDTITE		397
7 e	2418		2475
Qy	398 RYVLDISAEPWLTFGV		422
Q	: : : : : : : :		2535
ογ S	Qy 423 KLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGQIHIA 		467
3	0000		\$6C 7
ογ Op	QY 468PMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSV Db . 2595 APDPLSFTADOSYPTSCCAPREPGPAYTAREARGVPAESTRPAPGPRD		524 2644
ò	525 DYAELNPSIIVRAGSFDPSSOPNDRHVAFSTDG	GKNWFOG	564
, <u>a</u>	: :		2694
δ	265		612
QQ	2695 TQASATTQGWRPPAL		2753
οy	613 PANAQIRSDRVNPKTFYALSNGTFY		672
Dp	2754 PALPIDPVPPPVATGTVLPGGENRRPPLTSGPAPTPPRV	PVGGPORR	2800
Qy Dp	QY 673 LMLAASSGLYHSTNG-GSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGT		722 2860
οy	Oy 723 IGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDIGG		782
QQ	Db 2861 SGPPPTYLTLEG-	Λ9	2874
δy	QY 783 APSGSPSPSV-SPSASPSLSPSPSPSSPSSPSSPSSSPSSPSSPS		825
qq	2875 APGGPVSRRPTTRQPVA		2934
οy	826		885
QQ	2935		2987
οy	886 DNQIKPGLQVVNTGSSSVDLSTVTVRYWF		931
g	2988 EYPAVRFRAPQNRPSVPASASSINPRIGSSLSGVSSW-	ASSLALHID 3	3033
οy	932 IGCGNIRASFGS		
qq	Db 3034ATPPPVSLLQ 3043		

Search completed: July 2, 2002, 09:17:53 Job time: 361 sec

Perfect score:

Title:

Sequence:

,

OM protein

Run on:

Scoring table:

Searched:

Minimum DB seq Maximum DB seq

..

Database

```
P74375 synechocyst P74375 synechocyst P73032 synechocyst P3032 chlamydomon P91365 caenorhabdi P5146 epiphyas po P5146 epiphyas po P5146 epiphyas po P5146 epiphyas po P5148 postumis P5146 actinomyces P5148 arthrobacte P51318 clostridium P51475 paramecium P51475 streptomyce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ogibt9 turkey herp
Ogy076 leishmania
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=A3(2);

X MEDLINE=37000351;

X Redenbach M., Kleser H.M., Denapaite D., Eichner A., Cullum J.,

R Redenbach M., Kleser H.M., Denapaite D., Eichner A., Cullum J.,

R Redenbach M., Kleser H.M., Denapaite D., Eichner A., Cullum J.,

R Redenbach M., Kleser H.M., Denapaite D., Eichner A., Cullum J.,

R Rinashi H., Hopwood D.A.;

The M B Mb Streptomyces coelicolor A3(2) chromosome.";

Mol. Microbiol. 21:77-96(1996).

R EMBL; ALO3155; CAA20642.1; -.

R HSSP; P07986; LEXG.

R InterPro; IPR002860; BNR.

R InterPro; IPR002860; BNR.

R Ffam; PF002012; BNR; 10.

R Pfam; PF002012; BNR; 10.

R Pfam; PF00553; CBD_2; 1.

O SEQUENCE 890 AA; 93252 MW; D2C58695A4B56E84 CRC64;
                           09f2b0 thiobacillu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         29kyf3 streptomyce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
NCBI_TaxID=1902;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN=A3(2);
Seeger K.J., Harris D.;
Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Parkhill J., Barrell B.G., Rajandream M.A.;
Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               890 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ALIGNMENTS
                                                                                                                   Q39492
P91365
Q9AQG7
                                                                                                                                                                                                                                       Q9X3P6
Q9FPQ6
Q94F92
Q912M3
Q95L89
                                                                                                                                                                                                                                                                                                                                                                                 Q59164
Q9L448
Q96U11
Q46392
Q9N1P0
Q9L3J8
Q9E343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q84565
Q9AHP2
Q9F2X3
Q9IBT9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
                                                                       P73032
Q9RPL0
                                                                                                                                                                                           Q9RFX6
Q91GH4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                               707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2
10
                                                                                                                                                                                           100
100
100
100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-NOV-1998 (TrEMBLrel. 08, 01-NOV-1998 (TrEMBLrel. 08, 01-JUN-2001 (TrEMBLrel. 17, PUTATIVE SECRETED CELLULASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (TremBLrel. (TremBLrel. (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Streptomyces coelicolor.
                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN=A3(2);
                                                                                                                                                                                                                                       276
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
265.5
26
320.5
318.5
308.5
301.5
297.2
290.5
279
279
279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               086727
  71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

71188

7188

7188

7188

7188

7188

7188

7188

7188

718
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09wel thermotoga
059x91 caldicellul
09x35 caldicellul
09x35 caldicellul
09x35 caldicellul
09x311 anaerocellu
039620 chlamydomon
09x411 bacillus sp
044562 actinomyces
09ag4 caldicellul
09rx5 caldicellul
09rx5 caldicellul
                                                                                                                                                           2, 2002, 09:11:52; Search time 91.42 Seconds (Without alignments) 1810.940 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   O994h0 caldiceliul
O97kk0 clostridium
O994t8 agaricus bi
O74170 aspergillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  086727 streptomyce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                   5135
1 MDRSENIRLIMRSRRLVSLL......RASFGSVNPATPTADTYLQX 957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 562222
                      GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                               562222 seqs, 172994929 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUMMARIES
                                                                                                                   protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9AQH0
Q97KK0
Q9P4T8
O74170
Q9WYE1
O52374
Q95374
Q95311
Q3630
                                                                                                                                                                                                                                                                                                                                                                         Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sp_rodent:*
sp_virus:*
sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q92411
Q44562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q9AQG4
Q9RFX5
Q24820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  086727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_fungi:*
sp_human:*
sp_invertebrate:*
sp_mammal:*
sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sp_organelle:*
sp_phage:*
sp_plant:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   £
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sp_archea:*
sp_bacteria:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Post-processing: Minimum Match 0%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3
3
2
2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sp_rvirus:*
                                                                                                                                                                                                                                                           US-09-917-376-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SPTREMBL_19:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1770
921
1711
473
473
997
901
1751
930
                                                                                                                                                                                                                                                                                                                                                           BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10:
```

2436 2420.5 2037 1784 1740

1161

431.5 409.5 402.5 398.5 372.5 346.5 325.5 325.5

Score

Result ٠ ي us-09-917-376-1.rspt

```
19;
                                                                                                                                                                                                                                                                                                                                                                                                                                              909
                                                                                                                                 188
                                                                                                                                                                                                                                                                                                         546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               237
                                                                                                                                                                                                                      307
                                                                                                                                                                                                                                           297
                                                                                                                                                                                                                                                                 367
                                                                                                                                                                                                                                                                                     354
                                                                                                                                                                                                                                                                                                                                                       486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       607 AASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVAA-GLPSSGAVGVMFHA 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   666 VPGKEGDLWLAASS-----GLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVV 720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDI 780
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    900
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          811
                                                                                     VFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVW 128
                                                                                                                                                                          189 FGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSS-SS 247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  706 AEIGGVRGIFRSTDAGATWTRVNDDAHQWGWTGAAITGDPRVYGRVYVATNGRGVIYGD- 764
                     Gaps
                                                                57
                                                                                                                                                                                                                                                                            118 AAVGTYTNDWDPTNGAVLRSADRGASWERADLPFKLGGNNPGRGWGERLAVDPHDNDVLY
                                                                                                                                                                                       GGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISW
                                                                                                                                                                                                                                                                                                                                                      DEAMAIDPFNSDRMLYGTGATLYATNDLTKW-DSGGQIHIAPMVKGLEETAVNDLISPPS
                                                                                                                                                                                                                                                                                                                                                                                                             547 NDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSW
                                                                                                                                                                                                                                                                                                                                                                                                                                                        11 MRSRRLVSLLAATASFAVAAALGVLPIAITASPA--HAATTQPYTWSNVAIGGGGFVDGI
                                                                                                                                129 AAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILY
                                                                                                                                                                                                                      LGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNT
                                                                                                                                                                                                                                                                                                                                                                                                  GAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          841 SASPSPSSSPSSSPSSSPSPTPSSSPVSGGVKVQYKNNDSAPGDNQIKPGLQVVNTGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----TGACEVIYTVTNQWPGGFQ -- ADVRLTNTGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S----SVDLS----TVTVRYWFTR--DGGSSTLVYNCDW------AAIGCGNIRA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SAWNGWSLDWSFPGGOEVTRMWNAEHTQAGTSVTARNVGWNAGVAPGASVGFGFTGSRSG
                     Indels 126;
Length 890;
Score 2436; DB 2;
Pred. No. 7.4e-115;
); Mismatches 248;
                    Conservative 129;
47.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---SFGSVNPATPTA 951
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TNAEPEGFAVAGRACPTA 889
          Similarity
                     475;
Query Match
           Best Loc
Matches
                                                                                                                                                                                                                     248
                                                                                     69
                                                                                                                                                                                               178
                                                                                                                                                                                                                                                                                    298
                                                                                                                                                                                                                                                                                                                                                      428
                                                                                                                                                                                                                                                                                                                                                                                                  487
                                                                                                                                                                                                                                                                                                                                                                                                                     475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              765
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          783
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               901
                                                               qq
                                                                                     ŏ
                                                                                                          qq
                                                                                                                              δý
                                                                                                                                                 qq
                                                                                                                                                                         οy
                                                                                                                                                                                          g
                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                    δy
                                                                                                                                                                                                                                                                                                                                                                       OD
                                                                                                                                                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                      δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
```

```
STRAIN-TOKTB.1;
MEDLINE-20171169; PubMed=10706665;
Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P.,
Bargquist P.L.;
"Multidomain and multifunctional glycosyl hydrolases from the extreme
thermophile Caldicallulosiruptor isolate Tok7B.1.";
Curr. Microbiol. 40:333-340(2000).
EMBL, AF078038; AAK06388.1;
Interpro; IPR002860; BNR.
Interpro; IPR002965; P_rich_extensn.
Fran, PP02012; BNR; PR002180; PLICH_EXTENSN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTG 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             488 APLISALGDLGGFTHADVTAVPSTJFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQPN 547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    14 RRLVSLLAATASFAVAAALGVLPIAITASPAHAAT---TQPYTWSNVAI-GGGGFVDGIV 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKV-W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAVGMYTNSWDPNDGAILRSSDQGA/fWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSSL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PDTIIFRSTDGGATWTRIWDWTSYFNRSLRYVLDISAEPWLTFG-VQPNPPVPSPKLGWM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEETAVNDLISPPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3C72B6ED22F3C614 CRC64;
                                                                                                       01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       47.1%; Score 2420.5; DB 2;
llarity 48.5%; Pred. No. 5.1e-114;
Conservative 151; Mismatches 304;
                                                                                                                                                                            GLYCOXL HYDROLASE 5 (FRAGMENT).
Caldicellulosiruptor sp. Tok78.1.
Bacteria; Firmicutes; Bacillus/Clostridium group;
Thermoanserobacter group; Caldicellulosiruptor.
                                                      996 AA
                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             996
108275 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRINTS; PR01217; PRICHEXTENSN ProDom; PD001947; CBD_3; 1.
                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             996
996 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity
Matches 462; Conserv
                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hydrolase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NON TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Match
                                                   о9аоно
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query
RESULT
09AQH0
                                                         DDT TED DDT TE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

15;

126

54

186

305

466

406 484

424

603 585 763

721

```
Eukaryota: Fung1; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Agaricaceae; Agaricus.
                                                                             722 TIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDIG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 127 VWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNI
                                                                                                                             SLGQASKTIFVGVAD-PNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATS
                                                                                                                                                                                                                                                                                                                                                                                                                               586 NKWIPCSGLPQGAKVRSDRVNPKKFYGFLNGKFYISTDAGATFTQSSQTGLPTKGK-GI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FHAVPCKEGDLWLA-ASSGLYHSTNGGSSWSAITGVSSAVNVGFCKSAPGSSYPAVFVVG
               SRRLVSLLAA----TASFAVAAALGVLPIAITASPAHAAT-TQPYTWSNVAIGGGGFVD
                                                              67 GIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNK
                                                                                                                                                                LYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSS
                                                                                                                                                                                                                                                                 NTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQI
                                                                                                                                                                                                                                                                                                                366 SWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFG-VQPNPPVPSPKL
                                                                                                                                                                                                                                                                                                                                                                  GWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEETAVNDLISP
                                                                                                                                                                                                                                                                                                                                                                                                                 485 PSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 545 QPNDRHVAFSTDGGKNWFQ-GSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF-QPVAAGLPSSGAVGVM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         908
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Agaricus bisporus (Common mushroom).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (TrEMBLrel. 15, C
(TrEMBLrel. 15, I
(TrEMBLrel. 19, I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CEL6 PROTEIN PRECURSOR. CEL6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    782 GAPSGSPSPSVS 793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1: :|:||;
764 GS-QPTPTPSVT 774
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-OCT-2000 (
01-OCT-2000 (
01-DEC-2001 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9P4T8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9P4T8
                                                                                                                                                                                                                                                                                                                                                                  425
                                     ~
                                                                                                                                                               187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9P4T8
                                                                                                                                                                                       q
                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                       Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                              ò
                                                                                  g
                                                                                                                δ
                                                                                                                                      q
                                                                                                                                                              δ
                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                 δy
                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                       임
                                                                                                                                                                                                                                                                                                                                                                                                                òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              16;
                                                                                       665
                                                                                                                                                               705
                                                                                                                                                                                       785
                                                                                                                                                                                                               762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mobiling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q., Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I., Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J., Bennett G.M., Koonin E.V., Smith D.R.; and Endert G.M., Koonin E.V., Smith D.R.; Genome sequence and comparative analysis of the solvent-producing bacterium (lostridum acetobutylicum."; J. Bacteriol. 183:4803-4838(2001).

EMBL, AEO07608; AAK78895.I; -. InterPro; IPR002866; BNR.
InterPro; IPR002048; EF-hand.
                                                  607 AASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF-QPVAAGLPSSGAVGVMFHA
                                                                                                                                                     PSSSPSSSSSSSPSPTPSSSPVSG-GVKVQYKNNDSAPGDNQIKPGLQVVNTGSSSVD
                                                                                                                                                                                                                                                                                                      DRHVAFSTDGGKNWFQG-SEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSW
                                                                                                                                      VPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGG
                                                                                                                                                                                       VTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDIGGAPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
Clostridium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 839;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
, SEVERAL ASP-BOXES AND DOCKERIN
                                                                                                                                                                                                                                                                                                                                                     905 LSTVTVRYWFTRDGGSSTLVYNCDWAAIGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Complete proteome.
SEQUENCE 839 AA; 90824 MW; 6C99A041CA9CF984 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 39.7%; Score 2037; DB 16; Best Local Similarity 49.5%; Pred. No. 7.9e-95; Matches 392; Conservative 118; Mismatches 252;
                                                                                                                                                                                                                                                                                                                                                                                                                               Æ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
MEDLINE=21359325; Pubmed=11466286;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF01344; Kelch; 1.
PROSITE; PS00018; EF_HAND; UNKNOWN_2
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                      Q97KKO;
01-CCT-2001 (TEMBLrel. 18, C
01-CCT-2001 (TEMBLrel. 18, L
01-DEC-2001 (TEMBLrel. 19, L
PROBABLY SECRETED SIALIDASE,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF02012; BNR; 9.
Pfam; PF00404; Dockerin_1; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Clostridium acetobutylicum
                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=1488;
                                                                                                                                                                                                                                                                                                                                                                                                                           Q97KK0
                                                                                                                                                                                                                                                                                                                                                                                                    ٣
                                                            532
                                                                                                                                     999
                                                                                                                                                              646
                                                                                                                                                                                     726
                                                                                                                                                                                                             90,
                                                                                                                                                                                                                                      98/
                                                                                                                                                                                                                                                              763
                                                                                                                                                                                                                                                                                       846
                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
Q97KK0
                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                          g
                                                                                                                                                                                     ò
                                                                                                                                                                                                             g
                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                      δ
                                     ð
                                                          g
                                                                                   οy
                                                                                                            g
                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                              셤
```

```
DVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQ
                                                                                                                             Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                              297;
                                                                                                                                                                                                                                                                                                                                                                                                                     33.9%; Score 1740; DB 3;
llarity 43.5%; Pred. No. 6.9e-80;
Conservative 139; Mismatches 297;
                                                                                             Ā
                       738 ATSTISVSTITITRSTITITRS-SGNGSP
            Created)
                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                            ProDom; PD001821; CBD_fungal; 1. SMART; SM00236; fCBD; 1. PROSITE; PS00562; CBD FINGARY
                                                                                                                                                                                                                                                                                                                                                                        PS00562; CBD_FUNGAL;
856 AA; 89820 MW;
                                                                                            PRELIMINARY;
                                                                                                                            01-NOV-1998 (TrEMBLrel. 01-JUN-2001 (TrEMBLrel.
                                                                                                                 01-NOV-1998 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
Matches 366; Conserv
                                                                                                                                                                                                            NCBI_TaxID-5053;
                                                                                                                                                   AVICELASE III.
                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                          074170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              504
                                                                     RESULT
074170
            ò
                                g
                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                    13;
       Morales Almora P., Thurston C.F.;
Morales Almora P., Thurston C.F.;
"Molecular analysis of the cellulolytic genes in Agaricus bisporus.";
Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AJ292929; CAC02964.1; -.
HSSP; P00725; 2CBH.
InterPro; IPR002860; BNR.
InterPro; IPR000284; CBD_fungal.
Pfam; PF02012; BNR; 8.
Pfam; PF00734; CBD_i. 1.
SMART; SM00236; CBD_FUNGAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   620
                                                                                                                                                                                                                                                                                                                                                                                                                                             317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   79 YVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVGMYTNSW 138
                                                                                                                                                                                                                                                                                                                  DPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLW 198
                                                                                                                                                                                                                                                                                                                                                                VWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTIIFRST 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPSP----KLGWMDEAMAID 434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CGTTWVLINDDQHQYG-NWGQAITGDHANLRRVYIGTNGRGIVYGDIGGAPSGSPSPSVS 793
                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEETAVNDLISPPSGAPLISAL 494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFS 554
                                                                                                                                                                                                                                                         20 LAATASFAVAAALGVLPIAITASPAHAATTQPYTWSNVAI-GGGGFVDGIVFNEGAPGIL 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MPSRVKFWITLTLGTFSLVY----GOAASOSYRWONVKIGGGGGFVPGIVFNPSEKGLA
                                                                                                                                                                                                                                                                                                                                                                                                RSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVG
                                                                                                                                                                                                                                                                                                                                                                                                                                             VADP-NNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              615 NAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   564 -VMVSRNQATFNIFYGASGKTFYVSTDNGKTFSAHGSLGSATSVFDITVH--PSVSGDIW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TDGGKNWFQGSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                295 VYKYSIDTGAWTDITPVSGSDL --- FFGFGGLALDTQKNGTVMVAALNSWWPDGQIFRST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDD
                                                                                                                                                                                                                                    50;
                                                                                                                                                                                                             Length 806;
                                                                                                                                                                                                                                  277; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     525 SDYGVTWNQHFGAPDNVQGGKVAISADADIILWRTNGNG------
                                                                                                                                                             CEL6.
7C67E382E62FF341 CRC64;
                                                                                                                                                                                                            Score 1784; DB 3;
                                                                                                                                                                                                                        Pred. No. 4e-82;
                                                                                                                                                                                                                       Best Local Similarity 43.7%; pred. No. 4e-85
Matches 355; Conservative 131; Mismatches
                                                                                                                                                    POTENTIAL.
                                                                                                                                                  23 PC
806 CE
84779 MW;
                                                                                                                                                                                                            34.78;
43.78;
                                                                                                                                                              24 8
806 AA;
                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                             Query Match
                                                                                                                                          Signal.
SIGNAL
                                                                                                                                                                                                                                                                                                                                                  139
                                                                                                                                                                                                                                                                                                                                                                                                                                            259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             681
8
                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                       Dp
                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       63
Aspergillus aculeatus.
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                64 RLN-SDDTWIPLMDWVGNDTWHDW---GIDALAIDPVDTDRVYVAVGMYTNEWDPNVGSI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30 AALGVLPIAITASPAHAATTQPYTWSNVAI-GGGGFVDGIVFNEGAPGILYVRTDIGGMY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           89 RWDAANGRWIPLLDWVG---WNNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VFWSRDGGATWQAVPGAPT-GFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          206 TWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               325 SGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GTGATLYATNDLTKWDSGQIHIAPMYKGLEETAVNDLISPPSGAPLISALGDLGGFTHA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ||:| ||: | ||: | || ||: | 355 PIWEWNGYPSINYYSYDISNAPWIQDTTSTDQFPV---RVGMWVEALAIDPFDSNHWLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                385 RIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNP-PVPSPKLGWMDEAMAIDPFNSDRMLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 856;
                                                                                                                 Arai M., Takada G., Kawaguchi T., Sumitani J.;
Arai M., Takada G., Kawaguchi T., Sumitani J.;
Avicelase III from Aspergillus aculeatus.";
Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
EMBL; AB01551; BAA29031.1; -.
HSSP; P00725; 2CBH.
InterPro; IPR002860; BNR.
InterPro; IPR00254; CBD_fungal.
Pfam; PF02015, BNR; 7.
Pfam; PF02015, BNR; 7.
                                                                                                                                                                                                                                                                                                                                                                                                                1.
BE085983AF60ED76 CRC64;
```

18;

328

368

562

```
329 TRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWD 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              671 GDLWLAAS-SGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGA 729
                                                                                                                                                                                                       179 KVDSFPST-----SVTFVLFDEKSGEKGSPTPRIFVGCSEPKG-IFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GATWQITPLP---FKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    563 Q-GSEPGGVTTGG--TVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPA---NA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               120 GETWTIVNLDKYGIKVGGNEDGRNAGERLQVDPNFSSVLFMGT-TKYGLWKSEDFGKNWK
                                                                                                                                                                  QMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFW
                                                                                                                                                                                                                                                                                                                    269 SRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTW
                                                                                                                                                                                                                                                                                                                                                                                             220 TEDGGTTWNVLPNLPNDLIPLRGKIH--DGILYVTLSNALGPNGATRGAVMKYVIADQKW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    278 YDVTPMKGD-----FGYCGIDVQE---NVVIVSTLDRWYPHDEIFISLNGGETWRPLLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 389 WTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LYATNDLTKWDSGGQIHIAPWVK-----GLEETAVNDLISPPSGAPLISALGDLGGFTH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             503 ADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              617 QIRSDRVNPKTFYAL--SNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHA----VPGKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-RT69B.1;
Morris D.D., Gibbs M.D., Ford M., Thomas J., Bergquist P.L.;
"Family 10 and 11 xylanase genes from Caldicellulosiruptor sp.
Rt69B.1.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       730 YRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDI 780
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
EMBL; AF036924; AAB95326.1; -.
HSSP; Q06851; INBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bacteria, Firmicutes, Bacillus/Clostridium group, Thermoanaerobacter group, Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 1779
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001956; CBD_3.
InterPro; IPR001305; CBD_6.
InterPro; IPR001000; Glyco_hydro_10.
Pfam; PF00942; CBD_3; 3.
Pfam; PF02018; CBD_6; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Caldicellulosiruptor sp. Rt69B.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FAMILY 10 XYLANASE (EC 3.2.1.8).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PF00331; Glyco_hydro_10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (TrEMBLrel. (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=70295;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-JUN-1998
01-JUN-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       052374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   449
        152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HAD DE RELLE BERNER BER
                                                                                     q
                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEGUENCE FROM N.A.
SEGUENCE FROM N.A.
SEGUENCE FROM N.A.
STRAIN-MSB8 / DSM 3109;
MEDLINE-99287316; PubMed=10360571;
Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
"Evidence for lateral gene transfer between Archaea and Bacteria from
genome sequence of Thermotoga maritima.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               792
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AANGRWIPLLDWVGWNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQ 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  32 LGVLPIAITASPAHAATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWD 91
:| | :| | | :| | : | | : | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
                                                                                                                                                                                              DYAASTSTGTGAVALSADGDTVLLMSSTSGALVSKSQG---TLTAVSSLPSGAVIASDKS
                                                                                                                                                                                                                                                                            NPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYH
                                                                                                                                                                                                                                                                                                                            NDDQHQYGNWGQA--ITGDHANLRRVYIGTNGRGIVYGDIGGAPSG-----SPSPSV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STNGGSSWSAI-TGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 707;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     79496 MW; 8CD8743CBDA6A99F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22.6%; Score 1161; DB 16; 34.6%; Pred. No. 7.2e-51;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative 126; Mismatches 280;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bacteria; Thermotogales; Thermotoga.
NCBI_TaxID=2336;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1999 (TrEMBLrel. 12, 01-NOV-1999 (TrEMBLrel. 12, 01-JUN-2001 (TrEMBLrel. 17,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; AE001712; AAD35393.1;
TIGR; TM0305; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002860; BNR.
Pfam; PF02012; BNR; 9.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ENDOGLUCANASE, PUTATIVE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Thermotoga maritima
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Complete proteome.
SEQUENCE 707 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     854
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Sim
Matches 267;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -: Y
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9WYE1
Q9WYE1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                  527
                                                                                                                                                                                                                                                                                                                                                     584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        743
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    761
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92
                                                                                                                                                                                                                                                                        624
                                                                                                                                                                                                                                                                                                                                                                                                                                  684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 853
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09WYE1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCOORDING TO THE STATE OF THE 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                        a
                                                                                                                     ŏ
                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
```

Db 827 TKEIRLSIQKSSGSYNQSNDYSVRSANSYIENEKVTGYIDDVLVWGREPGRNAQIKVWYA 886 Qy 720 VGTIGGVTGAYRSDDCGTTWULINDDQHQYGNWGQATTGDHANLR 764	Oy 765 RVYIGTNGRGIVYGDIGGA	RESULT 8 O9X3P5 ID 09X3P5 O9X3P5 D 09X3P5 D 01X3P5 DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update) DT 01-NOV-1999 (TEMBLrel. 19, Last sequence update) DE XYNA. Caldicellulosiruptor sp. Tok7B.1. CALLINGTON NA. RECOURTE FROM N.A. RECOURTE FROM N.A. REQUISITE TAXID=80339; RN 11 RN 111-TAXID=80339; RN 61D58 M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P., RA BETQUIST P.L.; RA BETQUIST P.L.; RA HULLIGOMAIN and multifunctional glycosyl hydrolases from the extreme Tr thermophile Caldicellulosiruptor isolate Tok7B.1."; RL Curr. Microbil. 40:333-340(2000). DR REDEL, RAPOT9377; AAD30363.1; DR REPROSE PROFOUNCE (SPC.) DR REPROSE PROFOUNCE (SPC.) DR REPROSE PROFOUNCE (SPC.) DR REPROSE PROFOUNCE (SPC.) DR PROPOSE PROFOUNCE (SPC.) DR PRIMES: PROFOUNCE (SPC.) DR PROFOUNCE (SP	Query Match 8.0%; Score 409; DB 2; Length 1770; Best Local Similarity 20.0%; Pred. No. 1.38-12; Matches 263; Conservative 163; Mismatches 386; Indels 504; Gaps Qy 14 RRLVSLLAATASFAVAALGVLPIAITASPAHAATTQPYTWSNVAIGGGGFVDGIVFNEG 73 1:
DR PRINTS; PR00134; GLHYDRLASE10. DR PRODCH: PD001947; CLBD_3; 3. DR PROSTE; PS00591; GLXCOSYL_HYDROL_F10; 1. KW Xylan degradation; Hydrolase; Glycosidase. SQ SEQUENCE 1779 AA; 194304 MW; CE5269B6806B5CED CRC64;	Query Match Best Local Similarity 19.8%; Score 431.5; DB 2; Length 1779; Best Local Similarity 19.8%; Pred. No. 9.8e-14; Matches 252; Conservative 162; Mismatches 382; Indels 477; Gaps 53; Qy 65 VDGIVFNEGAPGILXVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAAD 120 ::	278 AVPGAPTGFIPHKGVEDPVNHVLYIA-TSNTGGPYDGSSGDVWFSV	

58;

```
δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                         8
                                                                                                     d
                                                                                                                                                      δy
                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         959 YVEIGFKSGAGVLAAGQSTKEIRLSIQKGSGSYNQSNDYSVRSATGYIENEKVTGYIDGA 1018
                                                                                                                                                                                                       TVKPDSVVAXSGKYS----LYVSGRTSNWHGAQIPVDTILEQGKVYKISVWVYQNSGST 262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APLISALGDLGGFTHADVTAVPSTI--FTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQ 545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | | : | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              171
                                                                                                                                                                                                                                                                                                                                                    PGAPTGFIPHKGVFDPVNHVLYIA--TSNTGG---PYDG--SSGDVWKFSV----TSGTW 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     :: | :: | 1:375 KMVLKHFNSITAGNEMKPSELLVDENTYNFSKADEFVNFATSNNIAIRGHTLVWHEQTPD 434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    663 KYAFWSLIEPTVIPANSTLPAPPAIQIPTPTPTPTPTVSATPTPA-PTASPVG-GSYW 720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LYHSTNGGSSWSA-----ITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGA--- 729
APGILY-----VRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNK 126
                                                                                                                                                                                                                                                                                                                                                                                                                                            TRISPVPSTDTAND-YFGYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIW 387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             435 WFFKDTNGNTLSKDALLSRLKQYIYTVVGRYKGKVYAWDVVNEAIDESQGDGFRRSNWYN 494
                           ADTLSYFAYGKSSIAVDMGN------AYNGKSSVRVS--NRSS 75
                                                                                                               DIQGVVWVAFDKSSSLGQASKTIFVGVADPN----NPVFWSR--DGGAT--WQA----V
                                                                                                                                                                                                                                                                                        -YRSDDCGTT------WVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGI-V
                                                                                       VWAAVGMYTNSWDPNDGA1LRSSDQGATWQIT-----PLPFKL-----GGNMPGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DWTSYPNRSLRY -----VLDIS--AEP---------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----YGTGATLYA------TNDLTKWDSGGQIHIAPMVKGLEETAVNDLISPPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              603 PRDLLIKQAMRYKELFDLFKKYNNVITSVTFWGLKDDYSWLSQNFGKSDYPLLFDENYKS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ---WFQGSEPGCVTTGGTVAA-----SADGSRFVWAPGDPGQPVVYAVGFGNSW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AASOGVPA------FYRSTBCVNPKTFYALSN-GT------FYRSTDGGVT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SVSVTSSIDPAYIDVKFVKLGANAGGADYYVEIGFKSGAGVLAAGQSTKEIRLSIQKGSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SYNOSNDYSVRSATGYIENEKVTGYIDDVLVWGREPSRNAQIKVWYANGNLSSPTNVLNP
                                                                                                                                                                           GMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PND------STDGGKN-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----FQP-----VAAGLPSSGA-----VGVMFHAVPGKEGDLWLAASSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        YGDIG---GA------PSGS------
                                                                                                                                                                                                                                                                                                                                                                                             208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             777
                                                                                     127
                                                                                                                                                                           172
                                                                                                                                                                                                                  126
                                                                                                                                                                                                                                                              232
                                                                                                                                                                                                                                                                                                                                                                                                                                         329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      730
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                         g
                                                                                                                          g
                                                                                                                                                                        ò
                                                                                                                                                                                                              Op
                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
δ
                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                      òγ
```

```
11;
                                                                                         1132
808
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       694 IIGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLIND-DQHQYG-- 750
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            827
                                                                                                                                                                                                                                                                                                                                                                         XYNA.
Caldibacillus cellulovorans.
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Alicyclobacillus group; Caldibacillus.
NCBI_TaxID=74586;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
MEDLINE=20519266; PubMed=11065373;
Sunna A., Gibbs M.D., Bergquist P.L.;
Sunna A., Gibbs M.D., Bergquist P.L.;
A novel thermostable multidomain 1.4-beta-xylanase from
'Caldibacillus cellulovorans' and effect of its xylan-binding domain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                   -- NWGQAITGDHANLRRVYIGTNGRGIVYGDIGGAPSGSPSPSVSPSASPSLSPSPSSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VSGGVKVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                       ASPSPSSSPSPSSSPSSPSPTPSSSPVSG-GVKVQYKNNDSAPGDNQIKPGLQVVNTGS
                                                                                                                               SSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAIGCGNIRASFGSVNPATPTADTYLQ 956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 2; Length 921;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL.
BETA-1,4-XYLANASE XYNA.
N; C5DDD1A7F7567413 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        On enzyme activity.";

Microbiology 146:2947-2955(2000).

EMBL. AF200304, AaF61649.1; -.

EMSP. OOG851; INBC.

InterPro: IPR001305; CBD_3.

InterPro: IPR001305; CBD_6.

InterPro: IPR001000; dlyco_hydro_10.

InterPro: IPR001000; dlyco_hydro_10.

InterPro: IPR001000; dlyco_hydro_10.

Pfam: PF00118; CBD_6; 1.

Pram: PF00118; CBD_6; 1.

PRINTS; PR00134; Glyco_hydro_10; 1.

PRINTS; PR00134; Glyco_hydro_10; 1.

PRINTS; PR00134; CBD_8; 2.

ProDom: PD001947; CBD_3; 2.

Signal; Xylan degradation; Hydrolase; Glycosidase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       689 OTTWGPS----TRITLYRNG-VLVWGT---EPGGG-----
                                                                                                                                                                                                                                                                                                01-007-2000 (TrEMBLrel. 15, Created)
01-007-2000 (TrEMBLrel. 15, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
BETA-1,4-XYLANASE XYNA PRECURSOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1,3e-12;
has 66;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; Pred. No. 1.3e
56; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              7.8%; Score 402.5; 33.2%; Pred. No. 1.36
                                                                                                                                                                                                                                                                921
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WAAIGCGNIRASFGSVNPATPTADTYLQ 956
                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    102380 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  921
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity
nes 89; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      921 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                    09L8L8;
                                                                                                                                                                                                                                                                091818
                                                                                                                                                                                                                            σ
                                                       842
                                                                                                                                 901
                                                                                                                                                                    1133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 751
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   929
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                          RESULT
Q9L8L8
```

```
g
                                δλ
                                                            a
                                                                                           δ
                                                                                                                      οp
                                                                                                                                                      ŏ
                                                                                                                                                                                   g
                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     40;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        262 PNNPVFWSRDGGATWQA----VPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDGS--S 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SWW------PDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQP 415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         252
                                                                                                                                                                                                                                                                                                          C-TERMINAL
                                                                                                                                                                                                                                                                                                                                                                                       C '- SIMILARITY: THE N-TEMNINAL FARKI DELONGS TO CELLULASE FAMILY L
C (FAMILY 9 OF CIYCOSYL HYDROLASES)
C '- SIMILARITY: THE C-TERNINAL PART BELONGS TO CELLULASE FAMILY L
C (FAMILY 48 OF CIYCOSYL HYDROLASES).
R HSSP; P26221; ITAT.
R HSSP; P26221; ITAT.
R INTERPO: IPRO01956; CBD.3;
R INTERPO: IPRO01956; CBD.3;
R INTERPO: IPRO01701; G1yco_hydro_48.
R Ffam; PF00942; CBD.3; 3:
R Ffam; PF00759; G1yco_hydro_9.
R PRINTS; PR00844; GLHYDRLASE48.
R PRODOM; PD001947; CBD.3; 2:
R PRODOM; P0010947; CBD.3; 2:
R PROSITE; PS00592; GIYCOSYL_HYDROL_F9_1; 1.
R PROSITE; PS00592; GIYCOSYL_HYDROL_F9_2; 1.
R PROSITE; PS00593; GIYCOSYL_HYDROL_F9_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             316 GDVWKFS---VTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQI----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  74 WAVYEYKDAFVKSGQLEHI--LNQIEWVNDYF-----VKCHPSKYVYYYQVGDGSKDH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGJIHIAPMVKGLEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TAVNDLISPPSG------APLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                 Anaerocellum thermophilum.
Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
Anaerocellum group; Anaerocellum.
                                                                                                                                                                                                                                                                                       FUNCTION: THE N-TERMINAL DOWAIN OF CELA ENCODES FOR AN ENDOCLUCARASE ACTIVITY ON CARBOXYMETHYLCELLULOSE. THE C-TERMI DOMAIN PROBABLY ACT SYNERGISTICALLY TO HYDROLYZE CRYSTALLINE CELLULOSE (BY SIMILARITY).

CELLULOSE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                    LINKAGES IN CELLULOSE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Glycosidase; Signal; Repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           7.8%; Score 398.5; DB 2; Length 1711;
llarity 22.9%; Pred. No. 4.3e-12;
Conservative 116; Mismatches 267; Indels 281;
                                                                                                                                                                                                                                                                        (MAR-1997) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1711 AA; 189979 MW; E3E987CEB9CD0C21 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             125 AWWGPAEVMOMERPSFKVTQSSPGSTVVTE----TAASLAAASIVLK---
                                              01-MAY-1997 (TIEMBLIEL. 03, Created)
01-MAY-1997 (TIEMBLIEL. 03, Last sequence update)
01-DEC-2001 (TIEMBLIEL. 19, Last annotation update)
ENDOGLUCANNSE A (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE CELL.
                              PRT; 1711 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Cellulose degradation; Hydrolase;
                              PRELIMINARY;
                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 197; Conserv
                                                                                                                                                                                                                                                             Zverlov V.;
Submitted (A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NON_TER
SEQUENCE
                              P96311
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         476
RESULT
            P96311
                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
17;
                                                                                                 253 KWAHCWDDVHNGAAL---LLAKITGKDIYKQIIESHLDYWITGYNGERIKYTPKGLAWLD 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               876 QYKNNDSAPGDNQIKPGLQVVNTGSS; VDLSTVTVRYWFTRDGGSSTLVYNCDWAAIGCG 935
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                          ----NAQIRSD---
                                                                                                                                                                                368 PPKRPHHRTAHSSWADSQSIPSYHRHTLYGALVGGPGSDDSYTDDISNYVNNEVACDYNA
                                                                                                                                                                                                                                                             --GSEPGGVTTGGTVAASAD------GSEPGWAPGDPGQPVVYAVGFG--
                                                                                                                                                                                                                                     -----RTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFH
                                                                                                                                                                                                                                                                                                                                                                                                                           535 SKNIYYILVDFTGTLIYPGGQDKYKKB-----VQFRIAAPQNVQWDNSNDYSFQDIKGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    767 YIGT------NGRGIVYGDIGGAPSGSPSPSVSPSASPSLSPSPSSSPSPSPSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Woesener J.P., Molendijk A.J., van Egmond P., Klis F.M.,
Goodenough U.W., Haring M.A.;
Domain conservation in several volvocalean cell wall proteins.";
Plant Mol. Biol. 26:947-960(1994).
                                                                                                                                                                                                                                                                                                              AVPGKEGDLWLAASS-----GLYHSTNGGSSWSAITGVSSAVNVGFGKSAP----
                                                                                                                                                                                                                                                                                                                                                                                           --GSSYPAVFVVGTI--GGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota, Viridiplantae, Chlorophyta, Chlorophyceae, Volvocales, Chlamydomonadaceae, Chlamydomonas.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           93;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 473;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
  ----AELNPSIIVRAGSFDPSSOPNDRHVAFSTDG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL.
38506131FAA674A4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7.3%; Score 372.5; DB 10;
33.0%; Pred. No. 2e-11;
tive 36; Mismatches 139;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           473 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL.
                                                                                                                                                          ------NSWAASQGVPA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-CC-621;
MEDLINE-95093034; PubMed-8000067;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     936 NIRASFGSVNPATPTADTYLQ 956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        750 NVTFKFVKLSSSVSGADYYLE 770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                47532 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01,
01,
19,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 7.3% Best Local Similarity 33.0% Matches 132; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Chlamydomonas reinhardtii.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1996 (TrEMBLrel.
01-NOV-1996 (TrEMBLrel.
01-DEC-2001 (TrEMBLrel.
VSP-3 PROTEIN PRECURSOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             32 4
473 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=3055;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         039620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
                                                                                                                                                                                                                                                                                                              665
                                                                             564
                                                                                                                                                        604
                                                                                                                                                                                                                                   522
527
```

σ

47;

292

386

458 470 487 763

737 801 832

857

```
; ; | | | ; | 355 WGSLRYAANAAFISFVYSDWVSDPVKKSRY -- QNFATSQINYILGDNPRQSSYVVGYGQN 412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SSPSPTPSSSPVS-GGVKVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTR 916
                                                         NELWGQVGAGNTDHAWWGPAE------VMQMNRPSFKIDASCPGSDLAAETAAAL 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              647 GVKVSGQEPGVTPVAVPAAPAG--VQAVA-GNSQVALNWSASAGAVSYTVKRAEVSGGPY 703
                                NKVWAAVGMYTNS---WDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDP 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDIGGAPSGSPSPSVSPSAS 797
                                                                                                               244 SWTGYEDELAWGGAWLYLAINDNAYLSKALSAADRWSTSGGSA-----NWPYTW
                                                                                                                                                                                                                                                                                                                                                                                                                                  SPQHPHHRTAHSSWMNNEDI--PANHRHILYGAMVGGPNASDQYTDDIGDYVSNEVATDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TTVAAGVNGLTYTNTGLTNGKTYYYVVTAVNSAGESPASVQVSGMPQAATTVPGAVTLSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           764 TAG---NNQNMLSWTAASGASSYT------VQRAVAG--GTYTDVATGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     --TGYQSDIQ-GVVWV-------AFDKSSSSLGQASKTIFVGVADPNNPVFW
                                                                                                                                                                                                                        SRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTW
                                                                                                                                                                                                                                                                                   TRISPVPSTDTANDYF - - GYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRI
                                                                                                                                                                                                                                                                                                                    -----KYTPGGLAWLDQ
                                                                                                                                                                                                                                                                                                                                                                                                                 ----TKW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          471 NAGFTGALAKMNLLYGONHOPLANFPAPEVKGDEYFVEAAVRSSGSNYTEIRALLNNRSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NNDNILYFGAPSGKGLWRSTDSGATWS------QMTNFPDV--GTYIANPTDT----
                                                                                                                                                                                                                                                    TQG----WDS-----KHYG----AQILLARITSNLNMP-----EATKF-----
                                                                                                                                                                                                                                                                                                                                                 ---DWTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DSG------ETAVNDLISPPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         488 APLISALGDLGGFTH----ADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----SFDPSSQPNDR-HVAFSTDGGKN-WFQGSEPG--GVTTGG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -- TVAASADGSRFVWAPGDPGQPVVYAVGFGNS-----WAASQGVPANAQIRSD-RVNPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TFYALS-NGTFYRST--DGGVTFQPVAAGL-----PSSGAVGVMFHAVPGKEGDLWLAA
 383;
 Indels
294;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DGGSSTLVYNCDWAAIGCGNIRASFGSVNPATPTADTYLQ 956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   802 AVLNYNDT-----TALNGTSYSYRVAAVNASGQTL---
 Mismatches
                                                                                                                                                                                                                                                                                                                  323 ----IQSTERNLDYWTVGTNGGRV----
 Conservative 109;
                                                                                                                                                                                                                                                                                                                                                   387 W-----
 214;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              13
                                125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  833
 Matches
                                                              147
                                                                                           182
                                                                                                                          196
                                                                                                                                                          227
                                                                                                                                                                                                                        269
                                                                                                                                                                                                                                                     293
                                                                                                                                                                                                                                                                                     329
                                                                                                                                                                                                                                                                                                                                                                                                               426
                                                                                                                                                                                                                                                                                                                                                                                                                                             413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      832
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                917
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
Q44562
                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                              음
                                                                                           οy
                                                                                                                          셤
                                                                                                                                                                                     g
                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                ò
                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-21129642; PubMed-11234960;
Pastor F.L.J., Dujol X., Blanco A., Vidal T., Torres A.L., Diaz P.;
"Molecular cloning and characterization of a multidomain endoglucanase from paenibacillus y BP-23: evaluation of its performance in pulp
                                                                                                                                                547
                                                                                           131
                                                                                                                                                                                       GVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSA---PGSSYPA 716
                               84
                   || | :|:| : | | : | | :|| SAAGPYTVYTYANSQVIASTLRLSLV---ETNLKY--LTPGGLARKNPIQTTSQPSAAQL
                                                                                          85 AGADAYVILSNRYYSYWDTEKMGSSPNKV------RDLKDWVNAGGS----LVLLDG
                                                                                                                         FGNSWAASQGVPANAQIRSDRVNPKTFYALSNG--TFYRSTDGGVTFQPVAAGLPSSGAV
                                                                                                                                                                                                                     -----SRGESGLTGCTSGAVLFSSNPDKMTKAVT--ASAITWSVGKGAITFIGSSFAM
                                                                                                                                                                                                                                                    717 VFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIV
                                                                                                                                                                                                                                                                                                                777 YGDIGGAPS----GSPSPSVSPSASPSLSPSPSSSSPSSPSSSSPSSSSPSSSPS--SPSP
SALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQPN----
                                                            ---DRHVAFSTDGGKNW---FQGSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVG
                                                                                                                                                                                                                                                                                   -----YEDMS----GVAVT--LANFNRT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Trefining.";

L Appl. Microbiol. Biotechnol. 55:61-68(2001).

C -1- SIMILARITY: CONTAINS 2 FTBRONECTIN TYPE III-LIKE DOMAINS.

R EMBL: Ad135614; CAB3491.1; -.

R HSSP; P26221; ITF4.

R HSSP; P26221; ITF4.

R INGEPTO: IPRO01956; CBD.3.

R INGEPTO: IPRO01701; Glyco_hydro_9.

B InterPro: IPRO01701; Glyco_hydro_9.

DR InterPro: IPRO01701; Glyco_hydro_9.

P Fam; PF00942; CBD_3; 2.

DR Pfam; PF00941; fn3; 2.

DR Pfam; PF00041; fn3; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               F20CB119D6410629 CRC64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-MAY-1999 (TrEMBLrel. 10, Created)
01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
CELLULASE PRECURSOR (EC 3.2.1.4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacillus/Staphylococcus group; Bacillus. (11 PaxID-89769; 11)
                                                                                                                                                                                                                                                                                                                                                                                                 SPSPSRSPSPSASPSSSPS--PSSSPSSPSPTPSSSP 868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 350.5; DB 2;
Pred. No. 5.9e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROSITE: PSO059; GLYCOSYL_HYDROL_F9_1; 1. PROSITE; PS00698; GLYCOSYL_HYDROL_F9_2; 1. PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00759; Glyco_hydro_9; 1.
ProDom; PD001947; CBD_3; 1.
SMART; SM00060; FN3; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             997 AA; 106927 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Glycosidase; Hydrolase; Signal SiGNAL 1 34 PSEQUENCE 997 AA; 106927 www
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-BP-23
                                                                                                                                                                                                                                                                                   PHLKG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         092411
                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 12
Q924I1
                                                                                                                                                                                       099
                             30
                                                            548
                                                                                                                         502
                                                                                                                                                                                                                                                                                   242
                                                                                                                                                                                                                                                                                                                                                264
                                                                                                                                                                                                                                                                                                                                                                               831
                                                                                                                                                                                                                                                                                                                                                                                                              320
                                                                                                                                                                                                                     191
                                                                                          g
                                                                                                                                                        g
                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                            ò
                                                                                                                         ò
                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                ò
ò
```

```
A Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P.,
A Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P.,
Bergquist P.L.;
The comparison and multifunctional glycosyl hydrolases from the extreme
the curr. Microbiol. 40:333-340(2000).

Curr. Microbiol. 40:333-340(2000).

R EMBL, ARV0842; ARV08394.1; -.

R EMBL, ARV0842; ARV08394.1; -.

R InterPro; IPR001956; CBD_3.

R InterPro; IPR001956; CBD_3.

R InterPro; IPR001956; Lipcoln_cytFABP.

R InterPro; IPR001956; Lipcoln_cytFABP.

R InterPro; IPR00796; Lipcoln_cytFABP.

R PRINTS; PR01217; PRICHEXTENSN.

R PRINTS; PR01217; PRICHEXTENSN.

R PROSTTE; PS00592; GLYCOSYL_HYDROL_F9_1; 1.

R PROSTTE; PS0013; LIPCORALIN; UNKNOWN_L.

R PROSTTE; PS0013; LIPCORALIN; UNKNOWN_L.

R SEQUENCE 1751 AA; 192176 MW; 60178CBF3C00BE95 CRC64;
                                                                                                                                                                                                                                                                                                          GWTWTHRTITADITKDKPWTARFAASGQGIQIQHGPHAGRLVQQYTIRTAGGAVQAVSVY 488
                                                                                                                                                                                                                                                                                                                                                                                                                                           651 NFT------MNWLGEQCGQKPAEPSPAPSPAPSAAPTEKPAPSAAPSAEPTQ 698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              :|:|||: ||::||| ||:|| ||:|| ||: || ||: ||-||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| ||:|| 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SLLAATASFAVAAALGVLPIAITASPAHAA-------TTQP---YTWSNVAI 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      60 GGGGFVDGIVFNEGAPGILYVRTD-----IGGMYRW-----DAANGRWIP 99
                                                         AVGFGNSWAASQGVPANAQIRSDRVNPKTFYAL----SNGTFYR----STDGGVTF-Q
                                                                                      SDHGKTWQA--GTPIGTGMDENKVVELSDGSLMLNSRASDGSGFRKVAHSTDGGQTWSE
                                                                                                                                                                   PVA-AGLPSSGAVGVMFHAVPGKEGDI,MLAASSGLYHSTN------GGSSWSAITGVS
                                                                                                                                                                                                                    PVSDKNLPDSVDNAQIIRAFPNAAPDDPRAKVLLLSHSPNPRPWSRDRGTISMSCDDGAS
                                                                                                                                                                                                                                                                             SAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQH---QYGN-WGQ
                                                                                                                                                                                                                                                                                                                                                                                       AITGDHANLRRVYIGTNGRGIVYGDIGGAPSGSPSPSVSPSASPSLSPSPS--PSSSSPSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 813 SPSPSSS--PSSSPSPSPSPSPSPSPSRSPS--PSASPSPSSSSPSPSSSSPSPSPSPSSSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 1751;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        869 VSGGVK-----VQYKNNDSAPGDNJIKPGL-----QVVN---TGSSS 902
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            759 AAEPTQAPTVAPSVEPTQAPGAQPSSAPKPGATGRAPSVVNPKATGAAT 807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Caldicellulosiruptor sp. Tok7B.1.
Bacteria: Firmicutes; Bacillus/Clostridium group;
Thermoanaerobacter group; Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6.5%; Score 334; DB 2; Lularity 21.2%; Pred. No. 7.7e-09; Conservative 125; Mismatches 314;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=20171169; PubMed=10706665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity
Matches 219; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=80339;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9AQG4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18
                                                      599
                                                                                                            489
                                                                                                                                                                   648
                                                                                                                                                                                                                       547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9AQG4
                                                      ò
                                                                                                          g
                                                                                                                                                                   ò
                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                              D
                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHP 356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           84 I------GGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVG 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             133 MYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAP 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       193 SGKGL-WRSTDSGATWSQMTNFPDVGTYIANPT-----DTTGYQSDIQGVVWVAFDKS 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----SDKTINVAATESSFDDLGRQCHWG------GLKPGKGAVYNCKPL 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           357 NTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPN 416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SDAPNPHIVQRRSTD-------GGKT-----WSAPTYIHQGTE 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TGTSVDYAELNPSIIV-----HVAFSTDG 557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GKNWFQGSEPGGVTTG----GTVAASADGSRFVWAP--------GDPGQPV-VY 598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             14 ALLGSLPLAATGLIAAAPPAHAVPT-----SDGLADVTITQVNAPADGLYSVGD 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         30 AALGVLPIAIT-----ASPAHAATTQPYTWSNVAIGGGGFVDGIVFNEGAPGI-LYVRTD 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   63 VMTFNITLINTSGEAHSYAPASTNLSGNVSKCRWRN-------VPPAGTTKT-DCTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              140 AGKALSTPETIKGATSPVKANSLRVESITPSSSQENYKLGDTVSYTVRVRSV-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SSSLGQASKTIFVGVA----DPNNPVFWSRDGCATWQAVPGAPTGFIPHKGVF---DPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        231 THTITOA----------DVDAGRWT-----------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            247 SITLTAT------GIDGATLQT------LTATGNPINVVGDHPQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      417 -PPVPSPKLGW-----MDEAM-----AIDPFNSDRMLYGTGATLYATNDLTKWDSGGQI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  465 HIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVF----T
                                                                                                                                                                                                                                                                                                                                                                                                                            "Complete nucleotide sequence of the Actinomyces viscosus T14V sialidase gene: Presence of a conserved repeating sequence among strains of Actinomyces sp.";
Extains of Actinomyces sp.";
EMBL: L06898: AAA1932.1; -.
HSSP: Q02834; IEUR.
Interpro: IPR002860; BNR.
Pfam: PF02012; BNR; S.
SEQUENCE 901 AA; 92861 MW; BCF6369DB6DF7E57 CRC64;
                                                                                                                                                                                                                    Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Actinomycineae; Actinomycetaceae; Actinomyces.NCBI_TaxID=1656;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      351;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 901;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 6.7%; Score 346.5; DB 2; Length 9 Best Local Similarity 22.5%; Pred. No. 8.4e-10; Matches 227; Conservative 101; Mismatches 330; Indels
                                                                                Last sequence update)
Last annotation update)
901 AA
                                                      Created)
PRT;
                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-93114861; PubMed-8418033;
                                                (TrEMBLrel. 01, C
(TrEMBLrel. 01, I
(TrEMBLrel. 19, I
PRELIMINARY;
                                                                                                                                                                                          Actinomyces viscosus.
                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                01-NOV-1996 (
01-NOV-1996 (
01-DEC-2001 (
                                                                                                                                 SIALIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       558
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

56;

```
575
                                                                                                                                                                                                                                                                                                                                                                                    519
                                                                                                                                                                                                                                                                                                                                                                                                                                                          708
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     813
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             841
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   246
                                                                                                                                                    489
                                                                                                                                                                                                                                                                                      413
                                                                                                                                                                                                                                                                                                                                   QPNPPVPSPKLGW----MDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPM 469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   634
                                                                                                                                                                                                                                     354
                                                                                                                                                                                                                                                                                                                                                                                                         VWGLEPGAATSTPAPTSTPTPTPTPTVTA-------TPTPTPTPTPTGSP--G 674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   694 ITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWG 753
                                                            351
-----YYIEGGKISGPYWD----KNRNIYYVLVDFSGTK----I
                                                                                                                                                                                                                                                                                                                                                                                                                                                         TGSGVKVLYKNNETSASTGSIRP------WFK-----IVNGG--SSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKE-GDLWLAASSGLYHSTNGGSSWSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 V-----SGADYY-LEVGF-SSGAGQLQ----PGKDTGDIQVRFNKNDWSNYNQADDWSW
                                                                                    ----ITPLPFKLGGNMPGR----GMGERLAVDPNNDNILYFGAPSG
                                                                                                                                                                                                         SNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTAN-DYF------GYSGLTIDRQ
                                                                                                                                                                                                                                                                                     355 HPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVL-DISAEPWLTFGV
                                                                                                                                                                                                                                                                                                                                                                                  VKGLEETAVNDLISP-----PSGAPLISALGDLGGFTHADVTAVPSTIFT---SPVFT
                                                                                                                                                                                                                                                                                                                                                                                                                                   TGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  580 ADGSRF---VW--APGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         709 VDLSRVKIRYWYTVDGDKPQSAVC-----DW------AQIGASNVT-FNFVKLSSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QAITGDHANLRRVYIGTNGRGIVYGDIGGAPSGSPSPSVSPSASPSLSPSPSSSSPSPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             O------TPAPTSTATPT
                                                                                                                                  KGLWRSTDSGATWSQMTNFPD----VGTYIANPTDTTGYQSDIQGVVW--VAFDKSSS
                                                                                                                                                                                  247 SLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIAT--
                                                                                                                                                                                                                                                            GNSQGP---NYTEVISYIYNRTGW----PPRVTDKLSFKYFIDLTELIQAGYS-----
                                                                                                                                                                                                                                                                                                                                                           576 YPGGEVEHKKQAQFKISVPQGYPWDPIN-DPSYKGLTSQLEKNKYIAAYDNNN----L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LQSMTN----YGENAKVTLY-----WG
                                    LLDWVGWNNWGYNGVVSIAADPINTNKVWAAVGM-YTNSWDPNDGAILRSSDQG-----
                                                           LAKITGKEL--YKGAVERNLDHWTDRITYTPKGMAYLTGW----GSLRYATTAAFLACVY
                                                                                                  930 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9RFX5;
01-MAY-2000 (TrEMBLrel. 13, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    : |:|||
----SVASSINPA 947
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CGNIRASFGSVNPA 947
                                                                                    ATWQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          O9RFX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            842
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     934
                                    100
                                                           862
                                                                                  153
                                                                                                          352
                                                                                                                                  195
                                                                                                                                                                                                           453
                                                                                                                                                                                                                                   305
                                                                                                                                                                                                                                                            490
                                                                                                                                                                                                                                                                                                           536
                                                                                                                                                                                                                                                                                                                                   414
                                                                                                                                                                                                                                                                                                                                                                                  470
                                                                                                                                                                                                                                                                                                                                                                                                          629
                                                                                                                                                                                                                                                                                                                                                                                                                                  520
                                                                                                                                                                                                                                                                                                                                                                                                                                                          675
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           801
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    754
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          825
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              939
                                                                                                                                                          401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9RFX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DA FE
                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            쉽
           셤
                                                          g
                                                                                 ò
                                                                                                        Op
                                                                                                                                                        g
                                                                                                                                                                                  ò
                                                                                                                                                                                                           g
                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                  ò
                                                                                                                                  Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
```

```
21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 231
                                                                                                                                                                      group; Caldibacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NWF-----QGSE----PGGVTTGGTVAASADG---SRFVW--APGDPGQPVVYAVGFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              664 HAVPGKEGDLWLAASSGLYHSTNGGSSWSAI--TGVSSAVNVGFGKSAP-GSSYPAVFVV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      :|| || || :|:|:|:|| :| 332 EFGDPREGGT------MITP------PPSGTTPTPTPTPTPTPSVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PSRSPSPSASPSSSSPSSSSPSSSPSTPSSSPVSGGVKVQYKNNDSAPGDNQIKPGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAIGCGNIRASFGSVNPATPTADT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          81 NWFGNLISNAGGRHREIIPDGKLCGPT--ASFDGMNQARTDWWTTRLOPGATITVRV---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAWAP------HPGTWY-----LYVTRDGWDPTQPL------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----KWSDLEPTPFSQVTNPPINSSGPDGAEY----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GTIGGVTGAYRSDDCGTTW-VLINDDQHQ---YGNWGQAITGD-HANLRRVYIGTNGRGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
MEDLINE=20120520; PubMed=10653733;
Sunna A., Gibbs M.D., Chin C.W.J., Nelson P.J., Bergquist P.L.;
Sunna A., Gibbs M.D., Chin C.W.J., Nelson P.J., Bergquist P.L.;
A gene encoding a novel multidomain beta-1,4-mannanse from
Caldibacillus cellulovorans and action of the recombinant enzyme
Kraft pulp.";
Appl. Environ. Microbiol. 66:664-670(2000).
EMBL; AF163837; AAF22274.1;
EMBL; AF163837; AAF22274.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             157;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL.
930 MULTIDOMAIN BETA-1,4-MANNANASE.
101576 MW; 0086638D54D1A2CC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length
01-MAY-2000 (TrEMBLrel. 13, Last sequence update) 01-DEC-2001 (TrEMBLrel. 19, Last annotation update) MULTIDOMAIN BETA-1,4-MANNANASE PRECURSOR.
                                                                            MANA.
Caldibacillus cellulovorans.
Caldibacillus firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Alicyclobacillus g
NCBI_TaxID=74586;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 UNKNOWN_1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6.3%; Score 325.5; DB; 25.8%; Pred. No. 9.9e-09;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             59; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IDR, 120. Chitin_bind_3.
InterPro; IPR001956; CBD_3.
InterPro; IPR001951; Chitin_bind_3.
InterPro; IPR001965; Chitch_extensn.
Pfam; PF00150; CBD_3; 2.
Pfam; PF00150; cellulase; 1.
Pfam; PF03067; Chitin_bind_3; 1.
Probom; PR01217; PRICHEXTENSN.
PROBOTITE; PS00659; GLXCOSYL_HYDROL_F5; UN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           34 5
930 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 109; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ХĽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           YIE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        834
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      954
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
     SETT DER KRANNER BETTER BETTER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
```

Search completed: July 2, 2002, 09:17:22

Job time: 330 sec

Run

```
ITTON Streptomyces coelicolor cosmid 5C7.

AL031515.

AL031515.

AL031515.

AL031515.

AL031515.

AL031516.

AL031516.

APP/CTP-binding protein; lyase; narG; narH; narI; narJ; nitrate reductase; oxidoreductase; RNA polymerase sigma factor; secreted cellulase; uvrA-like protein.

Streptomyces coelicolor A3(2).

Streptomyces coelicolor A3(2).

Bacteria; Firmicutes; Actinobacteridae;
Actinomycetales; Streptomycineae; Streptomycescence, Streptomyces.

CEL (bases 1 to 41906).

L Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                        ALO2000 Mindail Dan
X14717 Thermoprote
AC020697 Homo sapi
AC016142 Homo sapi
AC012149 Homo sapi
AC01712 Thermotog
AP00178 Homo sapi
AE00178 Homo sapi
AE00178 Homo sapi
AE00178 Homo sapi
AE00178 Homo sapi
                           AB01551 Asperdill
AE007608 Clostridi
AJ292929 Agaricus
AF200304 Caldibaci
D13755 Multicapsid
U75930 Orgyia pseu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BCT 07-SEP-1998
                                                                                                                                                                                                                                                                                      AF0386924 Caldicell
AF078737 Caldicell
E35099 Truncated c
AE007163 Mycobacte
AC021474 Homo sapi
AF210249 Streptomy
AF610249 Kanthomon
AE006990 Mycobacte
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2 (bases 1 to 41906)
Parkhill, J., Barrell, B.G. and Rajandream, M.A.
Direct Submission
Submitted (04-SEP-1998) Streptomyces coelicolor sequencing project, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CBIO 15A E-mail: barrell@aanger.ac.uk Cosmids supplied by Prof. David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
                                                                                                                                                                                                                                                                                                                                                                                                AE006990 Mycobacte
AL021897 Mycobacte
AF429315 Homo sapi
AL590867 Human DNA
                                                                                                                                                                                                                                              X14855 Thermoprote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Rattus no
Xylella f
Mycobacte
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 U10884 Bacteriopha
                                                                                                                                                                    Nostoc sp
Rattus no
                                                                                                                                                                                                                AY043265 Epiphyas
AF123462 Homo sapi
                                                                                                                                                                                                                                                            AP003821 Oryza sat
AL645725 Mus muscu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AE007164 Mycobacte
                                                                                                                                                                                                    Rattus no
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapi
                                                                                                                                        Caldibaci
                                                                                                                                                         Mus muscu
                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AE003939 > AL008967 N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AC011154
AC094068
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ALIGNMENTS
                                                                                                                                                                                                                                                                         AL645725
CASR69XYN2
AF078737
E35099
SC5C7
AF078038S1
                                                                                         NPORPT2
OPU75930
AR135147
                                                                                                                                                                                                                                                                                                                                                                   AF210249
AF069776
AE006990
MTV017
AF429315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AC068801
AC016142
AC012149
AE001712
AP001178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AC011154
AC094068
AE003939
MTV002
                                                                                                                                                                                                                                                                                                                                        AE007163
                                                                                                                                                                                                                                                                                                                                                                                                                                                AL590867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MTV023
                                                                                                                                      3.4 4567

3.1 183665

3.1 18364

2.9 19225

2.9 19824

2.9 18664

2.9 18604

2.1 186004

2.8 11007

2.8 11007

2.8 11007

2.7 16501

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.7 16502

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 1734

2.8 173
                                                                                                                                                                                                                                                           99427
216819
5437
11707
                                                                                                      3.6 131995
3.5 390
                                                                                                                                                                                                                                                                                                                                                                   77457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5513
47852
                                                                                                                                      96.4
91.4
90.2
86.4
883.6
83.6
81.6
81.6
                                                                                                                                                                                                                                                                                                                                                    748
512
402
375.6
327.6
131.6
103.6
103.6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
SC5C7/c
LOCUS
   \circ \circ \circ
                                                                                                                                                      O
                                                                                                                                                                                                                                                                                                                                         0000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OO
                                                                                                   2, 2002, 09:11:34 ; Search time 3473.69 Seconds (without alignments) 17283.705 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                 1 atggatcgttcggagaacat.......ggcggacacctacctgcagn 2869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Description
                                                                                                                                                                                                                                                                                                         3595312
              GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd
                                                                                                                                                                                                                                                                         1797656 seqs, 10463268293 residues
                                                                                                                                                                                                                                                                                                        of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                 Maximum Match 100%
Listing first 45 summaries
                                                                           - nucleic search, using sw model
                                                                                                                                                                                                                            IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              em_htq_other:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            em_htgo_inv:*
                                                                                                                                                                                                                                                                                                                                    length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              em_htg_hum:*
em_htg_inv:*
                                                                                                                                                                                                                                                                                                                                                                                   Post-processing: Minimum Match 08
                                                                                                                                                              US-09-917-376-2
2869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gb_ro:*
gb_sts:*
gb_un:*
gb_vi:*
em_ba:*
em_fun:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      em_or:*
em_ov:*
em_pat:*
em_ph:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    em_sts:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                           gb_ba:*
gb_htg:*
gb_in:*
gb_om:*
gb_ow:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gb_pat: *
gb_ph: *
gb_pl: *
gb_pr: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         em_mu:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            em_in:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              em_un:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                em_vi:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      *:ld_me
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   em_ro:*
                                                                                                                                                                                                                                                                                                                                                                                                                                            GenEmbl:*
                                                                                                        July
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Minimum DB seq
Maximum DB seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score
                                                                                                                                                                                   Perfect score:
                                                                                                                                                                                                                              Scoring table:
                                                                                                                                                                                                                                                                                                        Total number
                                                                           OM nucleic
                                                                                                                                                                                                 Sequence:
                                                                                                                                                                                                                                                                         Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                              Database
                                                                                                        on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      So.
                                                                                                                                                                   Title:
```

```
misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The more significant matches with motifs in the PKUSITE database are also included but some of these may be fortuitous. The length in codons is given for each CDS.

Usually the highest scoring match found by fasta -o is given for each CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(23:4768-4788|1994)) and the FramePlot program of Bibb et al., Gene 30:157-66(1984) as implemented at http://www.nih.go.jp/
jun/cgl-bin/framePlot.pl. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon. The may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid 577 lies between left and 485 on the Asel-A genomic restriction fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /db_xref="SPTREMBL:086698"
/translation="MQRGSDRMSVHRDDEMKHELQGLLRSGHPTRSEEWNDPEPAAED
                                                                                                                                                                                                                                                                                                                                by the BBSRC.

Details of S. coelicolor sequencing at the Sanger Centre are available on the World Wide Web.

(URL; http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are numbered using the following system eg SC7B7.01c. SC (S. coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /gene="SC5C7.02" uvrA-like protein, len: 796 aa; similar hote="SC5C7.02, uvrA-like protein, len: 796 aa; similar to Streptomyces peucetius TR:054827 (EMBL:1.76359) daunorubicin resistance protein drrc (764 aa), fasta scores; opt: 2051 z-score: 2702.2 E(): 0, 51.1% identity in 775 aa overlap, and to many uvrA proteins e.g. UVRA_NEIGO excinuclease ABC subunit A (950 aa), fasta
                                                                                                                                                                                                                                                                         Streptomyces coelicolor sequencing at The Sanger Centre is funded by the BBSRC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The more significant matches with motifs in the PROSITE database
                              3 (bases 1 to 41906)
Redenbach,M., Kleser,H.M., Denapaite,D., Eichner,A., Cullum,J.,
Kinashi,H. and Hopwood,D.A.
A set of ordered cosmids and a detailed genetic and physical map
for the 8 Mb Streptomyces coelicolor A3(2) chromosome
Mol. Microbiol. 21 (1), 77-96 (1996)
97000351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     complement(<1. .136)
/gene="SC5C7.01c"
/note="SC5C7.01c, unknown, partial CDS, len: >45 aa"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              complement(142. .146)
/note="possible RBS upstream of SC5C7.01c"
336. .339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /product-"hypothetical protein SC5C7.01c"
/protein_id-"CAA20613.1"
/db_xref-"GI:3559991"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /note="possible RBS upstream of SC5C7.02"
345. .2735

    .41906
/organism="Streptomyces coelicolor A3(2)"
/strain="A3(2)"

Colney, Norwich, Norfolk NR4 7UH, UK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /db_xref="taxon:100226"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           complement(1. .136)
/gene="SC5C7.01c"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /clone="cosmid 5C7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /gene="SC5C7.02"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /codon_start=1
                                                                                                                                                                                                                                                                     Notes:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                source
                                                                                                                                                                                         JOURNAL
MEDLINE
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gene
                                  REFERENCE
                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RBS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                  TITLE
```

```
2803. 3285
/gene="SC5C7.03"
2803. 3285
/gene="SC5C7.03"
2803. 3285
/gene="SC5C7.03"
/note="SC5C7.03"
/note="SC5.3"
/note="SC5C7.03"
/note="SC5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          te 2872. 3132
/gene="SC5C7.03"
/note="Pfam match to entry PF00165 HTH_2, Bacterial regulatory helix turn-helix proteins, araC family, score 71.40, E-value 1.9e-17"
/gene="SC5C7.04"
3291. 3707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WTVQVYAQSGFVDPDKPIREYTEKELRDFLYGEPVKVKVNGVNLTYEGLIPKIQKSFL
SKDKRAWQPHIRAFVERAVTFTCPECEGTRLSEGARSSKIKKISIDACAMETBLA
EWYDLTEPSVAPLLTARIDTLDSFVBIGLGYLSLDRPGTLSGGEAQRVKMIRHLGS
SLTDTTYVPDETYGLHPHDIQRANDLLRLRDKGNYVLVVEHKPEA.TAIADHYVDLG
PGAGTAGGTVCFEGTVEELRAADTVTGRHLDDRAVLKESVRKPAGALEIRDARTHNLQ
scores; opt: 412 z-score: 1697.0 E(): 0, 34.1% identity in 953 as overlap. Contains 2x PSO0017 ATP /GTP-binding site match f (P-loop) and 2x Pfam match to entry PF00005 ABC_tran, ABC_transporters, scores 42.40, E-value 1.7e-10 and 71.20, E-value 2.2e-17"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /traislation-"MSSAKRPGTPGPGSHVADSHDLIRVHGARENNLKDVSVDIPKRR
LTVETGVSGSGKSSLVFNIIAAESQLINETYSAFVQGFWFTARPEVDVLDGLITAI
LTVQQFWGADPRSTVGTATDVNAMLRILSRIGEPRIPSAYSFNTASGSAITV
ERGNKRAVRATFERTGGMCTHCEGRGTVSDIDLTQLYDDSKSLAGGAFTIPGWKSDSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QPAVLEYRFGGRDISEVLAMSVDQAEEFFGAGEARTPAAHKILQRLSDVGLGYLTLGQ
PLTTLSGGERQRLKLATHMGEKGGVYVLDEPTTGLHLADVEQLLGLLDRLVDAGKSVI
VIEHHQAVMAHADWIIDLGPGAGHDGGRVVFEGTPADLVADRSTLTGEHLAAYVGA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /translation="MTTLEDLARLRRARDLMDREYAEPLDVPALARVALMSAGHFSRS
FRAAYGETPYSYLMTRRIERAMALLRRGDMSVTEVCFAVGCTSLGSFSSRFTELVGES
PSAYRERDHADGAALPACVAKVYTRPVRNGEPGRSGTPAGKSESVRNGEAAPAPPPVA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GVDVDVPLGVLCVVTGVAGSGKSSLIHGSVPAGADVVSVDQSPIKGSRRSNPATYTGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LDPIRKAFAKANGVKPALFSANSEGACPTCNGAGVIYTDLAMMAGVATPCEDCEGKRF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3291. .3707
/gene="SCSC77.04"
/note="SCSC77.04"
LGUL_HAEIN lactoylglutathione lysse (EC 4.4.1.5) (135 aa),
fasta scores; opt: 124 z-score: 215.2 E(): 9.7e-05, 29.08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /note="PS00017 ATP/GTP-binding site motif A (P-loop)"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note="PS00017 ATP/GTP-binding site motif A (P-loop)"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /gene="SC5C7.02"
/note="Pfam match to entry PF00005 ABC_tran, ABC
transporters, score 42.40, E-value 1.7e-10"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1893. .2627
/gene="SC5C7.02"
/note="Pfam match to entry PF00005 ABC_tran, ABC
transporters, score 71.20, E-value 2.2e-17"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /product="putative transcriptional regulator"
/protein_id="CAA20615.1"
                                                                                                                                                                                                                                                                                                                                                                                                     /product="uvrA-like protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                  /protein_id="CAA20614.1"
/db_xref="GI:3559992"
/db_xref="SPTREMBL:086699"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /db_xref="SPTREMBL:086700"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /db_xref-"GI:3559993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1914. . . 1937
/gene="SC5C7.02"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    492. .515
/gene="SC5C7.02"
                                                                                                                                                                                                                                                                                                                                                        /transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /transl_table=11
                                                                                                                                                                                                                                                                                               /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /codon_start=1
```

m

```
33511 ATCAGCCCGCCGCGGAGGCGGACAC-----CTACTACGGCTTCAGCGCCTGACC 33461
                                                                                                                                                                                                                                                                                                                                                                                                                   GCCACGCGGACCCTCTACGTCGGGGTCGCGGACAAGGAGAACGCCGTCTACCGCTCGACC
                                                                                                                                                                                                 GGTGTGCTGGACGCGGAGAACGCCTACCTGTACCTCGCCTACAGCGACACCGGCGGCCCCG
                                                                                                                                                                                                                               tatgacgggagctccggcgacgtctggaaattctcggtgacctccgggacatggacgcga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    agtaa---gaccatttttgtgggggtggcggatcccaataatccggtcttctggagcaga
                                                                                                                                                                                                                                                                                      atcagcccggtaccttcgacggacacggccaacgactactttggttacagcggcctcact
                                                                                                                                                                                                                                                                                                                                                                                                                                                           1171 agttatcccaatcgaagcttgcgatatgtgcttgacatttcggcggagccttggctgacc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TCGTACCCGGACCGCGAGAACCGCTACACGATGGACGTCTCCTCGTCCCCCTGGCTCACC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gcaacaaatgatctcacgaagtgggactcc---ggcggccagattcatatcgcgccgatg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         gtcaaaagattggaggagacggcggtaaacgatctcatcagccgccgtctggcgccccg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       aatccgtcgatcatcgttcgcgctggaagtttcgatccatcgagccaaccgaacgacagg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            cacgtcgcgttctcgacagacggcgacagaactggttccaaggcagcgaacctggcggg
                                                                                                                 ggcgtctttgacccggtcaaccacgtgctctatattgccaccagcaatacgggtggtccg
                                                                                                                                                                                                                                                                                                                                                                                                      accataatctttcggagcaccgacggcggtgcgacgtggacgcggatctgggattggacg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            tcgacgatcttcacgtcaccggtgttcacgaccggcaccagcgtcgactatgcggaattg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gtgacgacggcggcaccgtcgccgcatcggccgacggctctcgtttcgtctgggctccc
                                                        754
                                                                                                                                                                                                                                                                                                                                            1051
                                                                                                                                                                                                                                                                                                                                                                                                    1111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1648
                                                                                                                                          33691
                                                                                                                                                                      871
                                                                                                                                                                                                 33631
                                                                                                                                                                                                                                                       33571
                                                                                                                                                                                                                                                                                      991
                                                                                                                                                                                                                                                                                                                                                                        33460
                                                                                                                                                                                                                                                                                                                                                                                                                                33400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       33340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               33280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              33160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       33100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              33040
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      32980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1708
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             32815
694
                                                                                 33751
                                                       δ
                                                                                                                                                                     Ω
                                                                                                                                                                                              g
                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                            δy
                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                             Dρ
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                              Óγ
                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                        /gene="SC5C7.05"
//gene="SC5C7.05, probable RNA polymerase sigma factor,
Inote="SC5C7.05, probable RNA polymerase sigma factors of the
sign subfamily e.g. TR:052313 (EMBL:AF036313) putative RNA
polymerase secondary sigma factor sigG (263 aa), fasta
scores; opt: 1288 z-score: 1540.5 E(): 0, 72.9% identity
in 251 aa overlap, and to RPOF.STRAU RNA polymerase
sigma-F factor (297 aa), fasta scores; opt: 713 z-score:
858.3 E(): 0, 42.6% identity in 312 aa overlap. Possible
alternative start 216 bp upstream, not supported by
FramePlot. Contains helix-turn-helix motif at aa 228-249
(Score 2104, 46.35 SD) and Pfam match to entry pr00140
                                                                  /db_xref="SPTREMBL:086701"
/translation="MDISLSQCFIAVDDHDRALAFYRDVLGLEVRNDVGYEGMRWVTL
                                                                                               GSPAQPDVDIVLEPPLADPNASAADKQAMADLLAKGHLRGIILATDDVDATFEHVRAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ώ
ώ
                                                                                                                                                                                                                                                                                                                                                                     sigma70, Sigma-70 factors, score 210.40, E-value 2.8e-59"
/codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CGGACCGAGAAGGACCTGGCCTACGCCCGACATCGGAGGCGCCTACCGTGGCAG 34232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            33812
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           33932
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                tacacctggagcaacgtggcgatcggggggggggcggctttgtcgacgggatcgtcttcaat 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       gaaggtgcaccgggaattctgtacgtgcggacggacatcggggggatgtatcgatgggat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACGGCGTGGTCGCTCTCCGACGCCGTCGATCCGGACCGGGTCTACGCGGCGGTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ggaatgggcgagcttgcggtggatccaaacaatgacaacattctgtatttcggcgcc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ggaatgtacaccaacagctgggacccaaacgacggagcgattctccgctcgtctgatcag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ggcgcaacgtggcaaataacgccctgccgttcaagcttggcggcaacatgcccggggcgt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CCCAGCGGCCACGGACTGTGGCGGTCCACCGACGCGGGCGTCACCTGGTCCGAGGTGACG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         33871 GCCTTCCCGAACCCCGGGAACTACGCGCAGGACCCGGAACGACGTCCGGCTACGCCTCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               aacggcgtcgtcagcattgcggcagacccgatcaatactaacaaggtatgggccgccgtc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       aactttccggacgtaggcacgtacattgcaaatcccactgacacgaccggctatcagagc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     64;
                                                                                                                                                                                                                                                                                                                                                                                                   /transl_table=11
/products"putative RNA polymerase sigma factor"
/db_xref="GI:3559995"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 41906;
                                                                                                                                        /note="possible RBS upstream of SC5C7.05"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                           GAEVLQEPVDQPYGVRDCAFRDPAGNMVRLHRPRKG"
3978. .3981
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mismatches 895;
                         /product="putative lyase"
/protein_id="CAA20616.1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 748;
Pred. No. 1
                                                                                                                                                               /gene="SC5C7.05"
3994. .4794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       26.1%;
59.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Simi
Matches 1429;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                34171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         34351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  34291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          34231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          34111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               34051
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        33991
                                                                                                                                                       gene
                                                                                                                          RBS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

us-09-917-376-2.rge

```
475 ccctgccgttcaagcttggcggcaacatgcccgggcgtggaatgggcgagcggcttgcg
                                                                                                                                                                                                    domain'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8; Mismatches 1168;
                                                                                                                                                                                                                                                                                             2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              17.8%; Score 512; DB 1; 52.4%; Pred. No. 3.2e-48;
                                                                                                                                                                                   274. .2580
/note="Region: endoglucanase
              2 (bases 1 to 3262)
Gibbs, M.D., Reeves, R.A., Farrington, G
Williams, D.P. and Bergquist, P.L.
Direct Submission
Submitted (27-OCT-1999) Biological Sc
Sydney, NSW 2109, Australia
Location/Qualifiers
                                                                                                                                                                                                            /evidence=experimental
274. .>3262
/note="ORF5; multidomain"
/codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        YGENAKVTLYVDGVLVWGQEPGGA"
                                                                                                                                                                                                                                                                                            /product="glycosyl hydro:
/protein_id="AAK06388.1"
                                                                                                                                                                 /db_xref-"taxon:80339"
274. .2580
                                                                                                                                                                                                                                                                                                                         /db_xref="GI:12743879
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note="PT-linker"
2791, .>3262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2791. .>3262
/note="type IIIb
529 c 794
                                                                                                                                                                                                                                                                             /transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .2790
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ø
 10706665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1056
                                                                                                                                                                                      misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Sim.
Matches 1358;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               415
 PUBMED
                              AUTHORS
                                                                             JOURNAL
                REFERENCE
                                                                                                        FEATURES
                                                             TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AF078038S1 3262 bp DNA linear BCT 11-FEB-2001 Caldicellulosiruptor sp. Tok7B.1 glycosyl hydrolase 5 gene, partial
                                                                             ACCGGCCTCCCGGCGGCGCGCATCGTCGAGTCGGAACCGGGTGAACCCCGCGGCGTTCTAC 32642
                                                                                                                                                                                                        32584 TCGGCCGCGACCGGCTGCCCGCTGACGCCTTCAAGGCGCTGCCGGGGG 32525
                                                                                                                                                                                                                                                               TTCGCAAGGCGGCCCCGGGAGCGTCGTACCAGACCCTCTTCACCAGCGCGGAGATCGGC 32345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               32224 CGGGTGTACGTGGCGACCAACGGCCGCGGGGTGATCTACGGCGACACCTCGGACACGGGC 32165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   32105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               32104 TACACGGTCACGAACCAGTGGCCGGGCGCTTCCAGGCGGACGTCCGCCTCACCAACACC 32045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2402
                                                                                                                                                                                                                                                                                                                                                                       2172
                                                                                                                         gocotatocaatggaacottotatogaagcacggacggcggcgtgacattocaaccggto 1947
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 caldicellulosiruptor sp. Tok7B.1.

Caldicellulosiruptor sp. Tok7B.1.

Bacteria; Firmicutes; Bacillus/Clostridium group;
Syntophomonadaceae; Caldicellulosiruptor.

E 1 (bases 1 to 3262)
Sibbs,M.D., Reeves,R.A., Farrington,G.K., Anderson,P., Williams,D.P. and Bergquist,P.L.
Multidomain and multifunctional glycosyl hydrolases from the extreme thermophile Caldicellulosiruptor isolate Tok7B.1

Curr. Microbiol. 40 (5), 333-340 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 the
                 GGGGGCACCGGTCCGGGTCCCGGACCCCACCCCACCCGCGCGCCTGAGGTGACG
ggcgatcccggtcagcctgtggtgtacgcagtcggatttggcaactcctgggctgcttcg
                                                             caaggtgttcccgccaatgcccagatccgctcagaccgggtgaatccaaagactttctat
                                                                                                                                         geggeeggtetteegageageagegtgeegteggtgteatgtteeaegeggtgeetggaaaa
                                                                                                                                                                                                                                              gaaggcgatctgtggctcgc-----tgcatcgagcgggctttaccactca
                                                                                                                                                                                                                                                                                                          accaatggcggcagcagttggtctgcaatcaccggcgtatcctccgcggtgaacgttggga
                                                                                                                                                                                                                                                                                                                                                                       tttggtaagtctgcgcccgggtcgtcatacccagccgtcttgtcgtcggcacgatcgga
                                                                                                                                                                                                                                                                                                                                                                                                                                   ggcgttacggggggggtaccgctccgacgactgtgggacgacctgggtactgatcaatgat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                GGGGTGCGCGCGATCTTCCGCTCCACCGACGCGGCGCGCGACCTGGACCCGGGTCAACGAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gaccagcaccaatacggaaattggggacaagcaatcaccggtgaccacgcgaatttacgg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2403 cccgagcccgagcccgagcagctcgccatcgccgtcgccgtcgccgagctcgagtccatc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     tggtgcgccgtccggatcgccgtctccgtcggtgagtccgtcggcttcgccgagcctgag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GCCACGTCGGCCTGGAACGGCTGGTCCCTCGACTGGTCCTTCCCCGGC 31997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2463 ctcgtcgccgtctccgtcgccgtcaccatcgccgagtccgtctcggtc 2510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           cgggtgtacataggcacgaacggccgtggaattgtatacggggacatt----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GI:12743874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AF078038
AF078038.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 2
AF078038S1
LOCUS
DEFINITION
                                                                                       32701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOURNAL
MEDLINE
1768
                                                           1828
                                                                                                                       1888
                                                                                                                                                                                   1948
                                                                                                                                                                                                                                              2008
                                                                                                                                                                                                                                                                           32524
                                                                                                                                                                                                                                                                                                          2053
                                                                                                                                                                                                                                                                                                                                                                      2113
                                                                                                                                                                                                                                                                                                                                                                                                                                  2173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                32344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            32044
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM
                             32755
                                                                                                                                                                                                                                                                                                                                       32464
                                                                                                                                                                                                                                                                                                                                                                                                    32404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          32284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  32164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEGMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE
                           g
                                                                                       a
                                                                                                                                               QQ
                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                           à
                                                                                                                     ò
                                                                                                                                                                                   ö
                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
/translation="wkrlishesillilnfliaefloggakadīgiavQpyvwknvki
EGGGGFITGIVFNPKEKNLVYVRTDIGGAYRSTDGGNTWTQLMDWVSFDDWNLLGVES
IATDPVDPNXVILACQGTYTNSWTDMNGAILRSTDEGDTFEITPLPFXLGGNMPXRNL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ó
                                                         Sciences, Macquarie University,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACAGATATGAATGGTGCTATTTTGCGTTCTACTGACGAAGGAGATACTTTTGAAATCACT 708
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             409 GGTGGAGGAGGTTTTATCACTGGAATAGTGTTTAACCCGAAAGAGAAAAATCTTGTTTAC 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gtgcggacggacatcggggggatgtatcgatgggatgccgccaacgggcggtggatccct 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      178 ggggggggggctttgtcgacgggatcgtcttcaatgaaggtgcaccgggaattctgtac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      cttctggattggggtgggatggaacaattgggggtacaacggcgtcgtcagcattgcggca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gacccgatcaatactaacaagg---tatgggccgccgtcggaatgtacaccaacagctgg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gacccaaacgacggagcgattctccgctcgtctgatcagggcgcaacgtggcaaataacg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GTTAGGACAGATATAGGAGGAGCTTACCGTAGTACAGACGGTGGCAATACATGGACACAA
Farrington, G.K., Anderson, P.,

    .3262
/organism="Caldicellulosiruptor sp. Tok7B.1"
/isolate="Tok7B.1"

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 3262;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    cellulose binding site"
g 871 t 12 others
```

Ŋ

QQ	709		
Oy Dp	535	gtggatccaaacaatgacaacattctgtatttcggcgccccgagcggcaaaggctctgg 594 	
Qy Dp	595 829	agaagcacagattccggcgcgacctggtcccagatgacgaactttccggacgtaggcacg 654 1 1 1 1 1 1 1 1 1	
Qy Db	655 889	tacattgcaaatcccactgacacgaccgctatcagagcgatattcaaggcgtcgtctgg 714 	
Qy Db	715	gtcgctttcgacaagtcttcgtcatcgctcgggcaagcgagtaagaccattttgtgggc 774	_
Oy Dp	775	gtggcggatcccaataatccggtcttctggagcagaggcgggggggg	~
Qy Db	835 1069	gtgccggtgcggaccggcttcatccgcacaagggcgtcttgacccggtcaaccac 894	
Oy Dp	895 1126	gtgctctatattgccaccagcaatacgggtggtccgtatgacgggagctccggcgacgtc 954	.0
Qy Db	955 1186	tggaaattctcggtgacctccgggacatggacgcgaatcagcccggtaccttcgacggac 1014 	10
Qy Dp	1015	acggccaacgactactttggttacagcgcctcactatcgaccgccagcacccgaacacg 1074 	- - 10
oy Q	1075	ataatggtggcaacccagatatcgtggtggccggacaccataatctttcggagcaccgac 1134 	10
Qy	1135 1357	ggcggtgcgacgtggacgcggatctgggattggacgagttatcccaatcgaagcttgcga 1194 	10
Qy Dp	1195	tatgtgcttgacatttcggcggagccttggctgaccttcggcgtacagccgaatcct 1251 	_ 10
Qy Dp	1252	cccgtacccagtccgaagctcggctggatggatgaagcgatggcaatcgatccgttcaac 1311 	_ 10
Qy Dp	1312	tctgatcggatgtctacggaacaggcggacgttgtacgcaacaatgatctcacgaag 1371 :	- 5
Qy Db	1372 1597	tgggactccggcggccagattcatatcgcgccgatggtcaaaggattggaggagacggcg 1431 	10
Qy Dp	1432 1657	gtaaacgatctcatcagccgccgtctggcgcccgctcatcagcgctctcggagacctc 1491 	0
Qy Db	1492 1717	ggcggcttcacccacgccgacgttactgccgtgccatcgacgatcttcacgtcaccggtg 1551 	
Qy	1552	ttcacgaccggcaccagcgtcgactatgcggaattgaatccgtcgatcatcgttcgcgct 1611 	_

2448 1908 2028 2088 agetegagtecatectegtegeegteteegtegeegteaceategeegagteegtetegg 2508 2568 2772 2817 1957 GCAGCTGCTGCAGATGGTAGTGCTGTTGTTTGGCCACCAAAAGGAGC-----GAAAGTA 2010 2011 TGCTATTCTACAGATAACGGAAACAAATGGGTTGAATGCGCAAATGTACCTTCAGAAGCT 2070 TGGGGTGAAATTGCTTATTC-----AAATATTTCGCCATCAGCAACTCCAACATCA 2592 2877 ggaagtttcgatccatcgagccaaccgaacgacaggcacgtcgcgttctcgacagacggc 1671 geogeateggeogaeggetetegtttegtetgggeteceggegateceggteageetgtg 1788 2362 ATTTATACTTATGCAAAATTAATGGAGTAAGAGGTATATTTAGATCTGATGACTGTGAT 2421 ATGGGGACTACTACTGATATTGATTTTGCTGAGTTAAATCCAAATTTTATGGTTCGTGTA 1836 ggcaagaactggttccaaggcagc---gaacctggcggggtgacgacgggcaccgtc 1897 GGAAAATCATGGTTTCAGGGAAATACTGAACCTCAAGGAACAAGCGAAGGTGGTACAGTT gigiacgcagicggattiggcaactccigggcigcticgcaaggigticccgccaaigcc 2071 ATTGTATATCTGACAGAGTAAATCCAAATAAGTTTTATGCATTTAAGAATGGCAAATTT tatcgaagcacggacggcgtgacattccaaccggtcgcggccggtcttccgagcagc geategageggetttaccacteaaccaatggeggeageagttggtetgeaateacegge accggtgaccacgcgaatttacggcgggtgtacataggcacgaacggccgtggaattgta teteegteaceateggegtegeegageeegtettegteaceg 2713 ACGCCGACGCCTACAGTGACGGTTACTATGACTCCGACACCGACACCAACACGCCG 2773 ACACCAACAGGGACACCTGGTAC------GGGAAGTGGTTTGAAGGTACTG cagatccgctcagaccgggtgaatccaaagactttctatgccctatccaatggaaccttc ggtgccgtcggtgtcatgttccacgcggtgcctggaaaagaaggcgatctgtggctcgct TTACCAATATCAGGAAATTTCAAAACTGTTCCAGGTATAGAAGGTGATATCTGGCTAGTT tacggggacattggtggtgcgccgtccggccgtccgtcggtgggtccgtcggct tegecgagectgagecegagecegagecegagetegecategeegtegeegt ACACCAAGTTTCACCCGACACCGACTCCAACAGCAACTATTACATCCACACCTGCACCA tcttcgtcgccgaggcccaacgccgtcgtcgtcgccggtgtcgggtggggtgaaggtgcag 2542 2389 1777 1612 1837 1672 1849 1969 2182 2029 2242 2149 2269 2482 2593 2449 2653 2509 2569 1729 1909 2329 2818 qq δý g a Qγ g Q ò QQ ŏ Op Qγ g ò Dp ογ Dp οy Db ò δ Ωp ŏ Dp δy qq δý g δ g ò δλ a ò οy

```
1043
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1020
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1223
 78
                                               138
                                                                                               198
                                                                                                                       326
                                                                                                                                             255
                                                                                                                                                                     386
                                                                                                                                                                                             315
                                                                                                                                                                                                                                          375
                                                                                                                                                                                                                                                                                          435
                                                                                                                                                                                                                                                                                                                                       495
                                                                                                                                                                                                                                                                                                                                                                                        555
                                                                                                                                                                                                                                                                                                                                                                                                                989
                                                                                                                                                                                                                                                                                                                                                                                                                                        609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                746
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              863
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              923
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  909
                         902
                                                                       997
                                                                                                                                                                                                                     446
                                                                                                                                                                                                                                                                   206
                                                                                                                                                                                                                                                                                                                  999
                                                                                                                                                                                                                                                                                                                                                                 626
                                                                                            g
                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                    g
                       Q
                                              g
                                                                      ò
                                                                                                                       δ
                                                                                                                                           P
                                                                                                                                                                   ò
                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                 QY
                                                                                                                                                                                                                                                                                       Dp
                                                                                                                                                                                                                                                                                                                                                                 óγ
                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                               òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       á
                                                                                                                                                                                                                                                                           Submitted (12-JUN-1998) Motoo Arai, Osaka Prefecture University, College of Agriculture; Gakuencho 1-1, Sakai, Osaka 599-8531, Japan (E-mail:motoo@biochem.osakafu-u.ac.jp, Tel:81-722-54-9465, Tel:81-722-54-9465)

2 (bases 1 to 2745)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  p mRNA linear PLN 20-JUN-1998
Avicelase III, complete cds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /protein_id="BAA29031.1"
/db_xref="G1:3242655"
/translation="MARSSLALLCAALLGKLADAAASQAYTWKNVTGGGGGFTPGIV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LYEGARSGHGLWKSTDYGATWSNVTSFTWTGTYFQDSSSTYTSDPVGIAWVTFDSTSG
SSGSATPRIFVGVADAGKSVFKSEDAGATWAWVSGEPQYGFLPHKGVLSPEEKTLYIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DGAAGLFKSEDAGTNWQVISDASHGFGSGSANVVNGDLQTYGRYFRGHERPGHLLRQS
QREPAGRHGDGDGDTTTSKTSTTVSTTLKTTTSSASTTSSASTTSKAS
STTTTKTTTTSTTTSSGTTATASAYAQCGGNGWTGATVCFTGYTCTYSNAFYSQCVPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13;
                                                                                                                                                                                         Aspergillus aculeatus cDNA to mRNA.

Aspergillus aculeatus

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

I (bases 1 to 2745)

Arai,M., Takada,G., Kawaguchi,T. and Sumitani,J.

Direct Submission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VAVGMYTNEWDPNVGSILRSTDQGDTWTETKLPFKVGGNMPGRGMGERLAVDPNKNSI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FNPSAKGVAYARTDIGGAYRLNSDDTWTPLMDWVGNDTWHDWGIDALATDPVDTDRVY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     cogetetgggagttetgcecategegataacggetteteetgcgcacgcggcgacgacte 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          77
                aataccgggtcgtcgtcggtggatttgtcgacggtgacggtgcggtactggttcacccgg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CCAGCCTGGCCTTGCTATGCGGGCCCTCCTGGGCAAGCTTGCAGACGCTGCGGCGTCGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            67;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 2745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 402; DB 8; Length 27
Pred. No. 4.9e-36;
); Mismatches 1020; Indels
                                                                                                                                                                                                                                                                                                                                      Arai, M., Takada, G., Kawaguchi, T. and Sumitani, J. Avicelase III from Aspergillus aculeatus Published Only in DataBase (1998) In press Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                     aculeatus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /product="Avicelase III"
                                                                                                                            Aspergillus aculeatus mRNA for i AB015511
AB015511.1 GI:3242654
Avicelase III.

    2745
    /organism="Aspergillus
    /db_xref="taxon:5053"

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                       'gene="avilII"
                                                                                                                                                                                                                                                                                                                                                                                                                                                              /gene="aviIII"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    14.0%;
llarity 52.7%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       O
                                                                                                                                                                                                                                                                                                                                                                                                                         .2578
                                                                                                                                                                                                                                                                                                                                                                                                                                                 .2578
                                              gatggtggctcgtc 2762
                                                                      GATGGTGACAAGCC 2951
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
Matches 1211; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     571
                                                                                                                                                                                                                                                                                                                                                                                        source
                                                                                                                                                       ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BASE COUNT
ORIGIN
                                                                                                                                          DEFINITION
                                                                                                                                                                                                                                         REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                           REFERENCE
AUTHORS
TITLE
JOURNAL
 2689
                                              2749
                                                                     2938
                                                                                                                                                                                                                                                                                                                                                                                                                            gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   149
                                                                                                        RESULT
AB015511
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                  CDS
                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                    g
ò
                      a
                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
```

```
1042
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      960 GTCTGTCCGTTGACCTCCAAGTTCCTGGCACCTTGATGGTGGCTGCACTCAACTGCTGGT 1019
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ggccggacaccataatctttcggagcaccgacggcggtgcgacgtggacgcggatctggg 1162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GGCCGGATGAGCTCATCTTCCGTAGCACTGACTCTGGAGCGACCTGGTCCCCTATTTGG 1079
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            attggacgagttatcccaatcgaagcttgcgatatgtgcttgacatttcggcggagcctt 1222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AGTGGAACGGTTATCCGAGCATCAATTACTACAGCTATGATGATATCAGCAACGCCCCTT 1139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ggctgaccttcggcgtacagccgaatcctcccgtacccagtccgaagctcggctggatgg 1282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             137
                                                                                                   254
                                                                                                                                                                                                                                                                                                                                                                                     cogcogtoggaatgtacaccaacagetgggacccaaacgacggagcgattetecgetegt 445
                                                                                                                                                                                                                                                                                                                                                                                                                505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             685
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           atcagagcgatattcaaggcgtcgtctgggtcgctttcgacaagtcttcgtcatcgctcg 745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          862
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         848
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CGGACCAGGGTGATACCTGGACTGAGACGAAGCTGCCGTTCAAGGTTGGTGGGAATATGC 434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CCGGTCGCGGAATGGGTGAGATTGGCGGTCGACCCGAACAAGAATAGCATTCTCTACT 494
ccgggcgtggaatgggcgagcggcttgcggtggatccaaacaatgacaacattctgtatt
                                                                             ctgatcaggggcgcaacgtggcaaataacgccctgccgttcaagcttggcggcaacatgc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  789 cocacaagegegretrerecergagaagaagacergracarerecaargege
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        gcctcactatcgaccgccagcacccgaacacgataatggtggcaacccagatatcgtggt
                                                                                                                                                                                                                                                                                    gggggtacaacggcgtcgtcagcattgcggcagacccgatcaatactaacaaggtatggg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  teggegeecegagegaaagggetetggagaageacagatteeggegegeetggteee
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           agatgacgaactttccggacgtaggcacgtacattgcaaatcccactgacacgaccggct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATGTCACCTCCTTCACGTGGACTGGAACGTATTTCCAGGATTCGAGCTCGAC-----CT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACACTTCGGACCCTGTGGGTATTGCTTGGGTGACCTTTGATTCGACCTCTGGTAGCTCGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ggcaagcgagtaagaccatttttgtgggcgtggcggatcccaataatccggtcttctgga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        cgcacaagggcgtctttgacccggtcaaccacgtgctctatattgccaccagcaatacgg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gtggtccgtatgacgggagctccggcgacgtctggaaattctcggtgacctccgggacat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               849 CTGGTCCTTACGATGGCACAAACGGAACAGTCCACAAGTACAACATCACCAGCGGAGTGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ggacgcgaatcagcccggtaccttcgacggacacggccaacgactactttggttacagcg
```

27-JUL-2001 the

```
Submitted (24-JUL-2001) GTC Sequencing Center Production, Finishing, and Bioinformatics teams, Genome Therapeutics Corp., 100 Beaver Street, Waltham, MA 02453-8443, USA Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /translation="mnkrivsmvaglslifttgfvthlsaaankaaasvqsqgqkwdn
akigaggyvpavifnktekdliyaktdbggayrwdkannkwipitdgfsDwtmlgces
iatdpidtnrvyiaaglytnDwqDenayilsSqDkGntwkRyQLPfkVgGnwpGRnwg
                                                                                                                                                                                                                                                                          1—(bases I to 12732)
Nolling,J., Breton,G., Omelchenko,M.V., Markarova,K.S., Zeng,Q., Gibson,R., Lee,H.M., Dubois,J., Qiu,D., Hitti,J., Wolf,Y.I., Tartusov,R.L., Sabathe,F., Doucettee-Stamm,L., Soucaille,P., Daly,M.J., Bennett,G.N., Koonin,E.V. and Smith,D.R. Genome Sequence and Comparative Analysis of the Solvent-Producing Jacterium Clostridium acetobutylicum
J. Bacterium Clostridium acetobutylicum
J. Bacterium (183 (16), 4823-4838 (2001)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /transl_table=11
/product="Probably secreted sialidase; several ASP-boxes
                                                                                                                                                                                                                                                             Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /transl_table=11
/product="Protein related to MIFH/DOPD protein family,
function in bacteria is unknown"
                                                                                                                              AE007608 12732 bp DNA linear BCT Clostridium acetobutylicum ATCC824 section 96 of 356 of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KDVTIDKKAADVNGDGTVNGRDLMTLRQYISGKITSFSSTK'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1. .12732_
/organism="Clostridium acetobutylicum"
/strain="ATCC 824"
/db_xref="taxon:1488"
                                                                                                                                                                                                                                                                                                                                                                                                                                                       2 (bases 1 to 12732)
Childress,D., Zeng,Q. and Smith,D.R.
Direct Submission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               and dockerin domain"
/protein_id="AAK78895.1"
/db_xref="GI:15023820"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 complement(2634. .2978)
                                                                                                                                                                                                                           Clostridium acetobutylicum. Clostridium acetobutylicum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                complement(2634.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /gene="CAC0919"
73. 2592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /gene="CAC0919"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /gene="CAC0920"
                                                                                                                                                              complete genome.
AE007608 AE001437
AE007608.1 GI:15023819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /codon_start=1
                                                  2244 GGCGCCACGGCGACGGCG 2261
                  ggatcgccgtctccgtcg
                                                                                                                                                                                                                                                                                                                                                                                                                                           11466286
                                                                                                                                                                                                            KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        source
                                                                                                                                               DEFINITION
                  2356
                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               gene
                                                                                                  RESULT AE007608
                                                                                                                                                                              ACCESSION
                                                                                                                                                                                                                                                                                                           AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JOURNAL
                                                                                                                                                                                                                                                                                           REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                         JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FEATURES
                                                                                                                                                                                             VERSION
                                                                                                                                                                                                                                                                                                                                                                           TITLE
                                               g
                    à
                                                                                                                                                              1462
                                                                                                                                                                                                                             cccgetcatcagegeteteggagaceteggeggetteaceaegeegaegttaetgeeg 1522
                                                                                                                                                                                                                                                                                                                          1493
                                                                                                                                                                                                                                                                                                                                                          aattgaatccgtcgatcatcgttcgcgctggaagtttcgatccatcgagccaaccgaacg 1642
                                                                                                                                                                                                                                                                                                                                                                           1643 acaggcacgtcgcgttctcgacagacggcggcaagaactggttccaaggcagcggaacctg 1702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       gcggggtgacgacggcgcaccgtcgccgcatcggccgacggctctcgtttcgtctggg 1762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  gcagcagttggtctgcaat -- caccggcgtatcctccgcggtgaacgtgggatttggta 2119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AGGCGTCCTCCACGGCCAGCTACGTCGTGATCTACGGGTTCTTCACCATCGACGCTGCCG 2063
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             cgttgtacgcaacaaatgatctcacgaagtgggactccggcggccagattcatatcgcgc 1402
                                                                                                                                                                                 CAACCAGCACCGGCACCGGAGCGGTGCTTTTCCGCCGATGGTGACACGGTGCTTCTCA 1658
 GGATCCA----GGACACGACTTCGACCGACTATTCCCGGTTCG--CGTCGGCTGGATGG 1193
                                                  ATCCGACCCTGGCTCTGTCGTCGAACTTCGGATCGACTTGGTATGCAGACTACGCCGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ctcccggcgatcccggtcagccgtgtgtgtagcagtcggatttggcaactcctgggctg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1884 CCATCGCCGGCGATGTCTGGGCGTCGACAAGGCCTCTGGCATTCCACGATTACG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 cgatggtcaaaggattggaggaggaggcggtaaacgatctcatcagcccgccgtctggcg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1659 TGTCCAGCACCTCGGGCGCGTTGGTGTCGAAGTCGCAGGG ------CACCTTGACTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    cttcgcaaggtgttcccgccaatgcccagatccgctcagaccgggtgaatccaaagactt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               cggtcgcggccggtcttccgagcagcggtgccgtcggtgtcatgttccacgcggtgcctg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -- ACCAAGACCGTGAGCCTGGGGTCCAGCACGTCAACGCCATCCGCGCGCATCCCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GTAGTACGTTCACGCAGATCGGCAGCGGCGTCACGGCGGGCTGGAGCTTTGGGTTTGGCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             agtotgogcocgggtcgtcatacccagccgtctttgtcgtcggcacgatcggaggcgtta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             cgggggggtaccgctccgacgactgtgggacgacctgggtactgatcaatgatgaccagc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACGGATTCGGCTCGGGTCCGCCAACGTGGTGAACGGCGACCTGCAGACCTACGGCCGCG
                                atgaagcgatggcaatcgatccgttcaactctgatcggatgctctacggaacaggcgcga
                                                                                                                                                                                                                                                                                          tgccatcgacgatcttcacgtcaccggtgttcacgaccggcaccagcgtcgactatgcgg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gaaaagaaaggcgatctgtggctcgctgcatcgagcgggctttaccactcaaccaatggcg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             accaatacggaaattgg.--ggacaagcaatcaccggtgaccacgcgaatttacggcggg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1599
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1943
1140
                                                              1194
                                                                                              1343
                                                                                                                                                                                           1314
                                                                                                                                                                                                                                                          1374
                                                                                                                                                                                                                                                                                          1523
                                                                                                                                                                                                                                                                                                                                                        1583
                                                                                                                                                                                                                                                                                                                                                                                     1494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1883
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1826
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2063
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1944
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2184
                                                                                                                                                             1403
                                                                                                                                                                                                                           1463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2297
                                 1283
                                                                                                                                                                                                                                                                                                                                                                                                                                                       1539
                                                                                                                           g
                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                            ò
                                                                                                                                                             ò
                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
a
                                ò
                                                            g
```

gene

CDS

gene

CDS

gene

CDS

```
RSIYLGEMSEGVYYNLALCMESIGELDDAIKWLNRALIINDONIITLNKKAAVLIKIK
KYEALLEVYNVUKIEVDNEEVYHFKAMLGEBGKKIERAFETISRGELGEGOWTLCY
DKAVLFEKOKEFEKALLGTEKALLLFDDENVLLIKKGDLLTYTKKIDEAKSYYDKILE
FHENNMEGVFAKANLCMLLGEHEDAEFFKEIIERLSGEESVLINSYYYRALNIKKWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                          /protein_id="Aak78902.1"
/bcarein_id="GI:15023827"
/transld=ion="MakKENEEREAKEYNFLIIKYPYDVKLNFFKANCLRDIED
YEKABELYEYIIDLDNDALEVYLMRARNRISLNKYEAAKEDLNTVIKLNPAYRKVIEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /product="Hypothetical protein, CF-2 family"
/protein_id="AAK78903.1"
/brotein_id="AAK78903.1"
/d_xref="GI:15023828"
/translation="WINSDERIBERACDDVIKDLKAVGEKVKETYNLIHLSNRHIEDG
IKGSAVDSITTEYNKIVKMVKEMQDNMGNDVAAVNNIVSSFDAEDKMIKTKFYR"
       AYLNKAALCLAMNKVDDAVSCYDKLIIKYPNEGKAYFGKGNVMFFYKDNIKKAIELYN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /transl_table=11
/product = "Uncharacterized small conserved protein, homolog of yfjA/yukE B.subtilis"
/protein_id="AAK78904.1"
/prodein_id="AAK78904.1"
/profession_id="AAK78904.1"
/prodein_id="AAK78904.1"
/profession_id="AAK78904.1"
/prodein_id="AAK78904.1"
/profession_id="AAK78904.1"
/prof
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /translation="MGSMQYFDNIAASWNVIRSEYFEERLKYKVLSKVSIKDKITADF
GCGTGFISLALARDANNYFGELDNSNMHIRELKEKSHDKKLMNIYLLKSSLTSVTLFDE
SIDAVFINMALHHIRDAEA AIKEWIRVLKKGGTVVISDVSEHNGKWAREEMFDEWLGF
SNEQIDWWLKNVGFKNIDIENTDLKCKGTSSRGEHTETSIFIAKAAK"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .;
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        154 tacacctggagcaacgtggcgatcggggggggggcggctttgtcgacgggatcgtcttcaat 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    214 gaaggtgcaccgggaattctgtacgtgcggacggacatcggggggatgtatcgatgggat 273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 375.6; DB 1; Length 12732; Pred. No. 2.8e-33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /product="SAM-dependent methyltransferase"
/protein_id="AAK78905.1"
/db_xref="GI:15023830"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                             /product="TPR-repeat-containing protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /transl_table=11
/product="Cystathionine gamma-synthase"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mismatches 1049;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEQLKEFLSGKSESCYR'
8379. .8675
                                                                                                                                                                                        /gene="CAC0926"
/gene="CAC0926"
7998. Rif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9712. .10335
/gene="CAC0929"
9712. .10335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /codon_start=1
/transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                /transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10353. .11498
/gene="CAC0930"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8379. .8675
/gene="CAC0927"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8713. .9012
/gene="CAC0928"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /gene="CAC0928"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /gene="CAC0929"
                                                                                                                                                                                                                                                                               7998. .8357
/gene="CAC0926"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /gene="CAC0927"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /gene="CAC0930"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0;
                                                                                                                                                                                                                                                                                                                                                        /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /codon_start=1
                                                                                                                                                                                 KTRKOKRLIRKO"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10353. .11498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13.1%;
50.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note="metB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Sim
Matches 1121;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         gene
                                                                                                                                                                                                                     gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         gene
                                                                                                                                                                                                                                                                                   CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /translation="motified@beflkkligyrrelhkhpelsmkeyettkrikrwlye
NDIKVLDFPIEVGVVAEIEGBHEGKTVAIRADIDALPIEEKTELPFSSVNKGIMHACG
NDIKVLDFPIEVGVVAEIEGBHEGKTVAIRADIDALPIEEKTELPFSSVNKGIMHACG
HNFPHTAALIGTAILLSKRKAEIYGRYRIIFQPGEETGKGSQI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /translation="MOSLKVDMRNGLLKINEAKQGLKSAVDKFSLKNKSRSIEIELKN
KREEKKAVNGLEGETRKINSNIDYMYKRFTDDVRKGASKLKTSGYTYRKYGLLNFGE
EVERVPVIGGIGVPIDAVESWVDRNKFVLKKTTTTILETGAIGVCIVTLPEAGITWKA
WDALSAAGGWASLDMIGTAAAAIYORKVERKSEEDSSGYDLFGOSIFGGAGWYSAKILG
VKNPSYWGSSGKTLYDRTSATVYLNVGHSVTSLGKDVEKLPALKKAEASSANKTLEA
ANNAEKIHDMAMATRNKINGINNAKEVVDLKKLKKNVGKTNFEPYKNGFYKDKNLL
KKDGNKLPOLGQDALLIQQDOLETFEKNTLANKNKYNKSIRDLNKNYRDIKASLYNIPA
GGTLYNTEDVQGSFHFNYFPGEIKPIFEEIKVVLGGER"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /product="Uncharacterized small conserved protein, homolog of yfja/yukE B.subtilis"
/protein_id="AAK78898.1"
/db_xref="G1:1505888.1"
/translation="WFDRAALECEDIKKSIEKSTTOLIDSWYGESCTAFEKTYLILDR KLNTYTEILQEYSKALYDISETYQNTDGDIAKAMSKAVNEASEQK"
                                /db_xref="G1:15023821"
/translation="MPFINSTVTVKLDKEKESTLKVELGKIIELLPGKSEDWLMVGFK
DNYPLYFKGEKKDKAAFIEVKIFGSADKASKNKLVSAISSLMENELSIPKDCIYITVD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /translation="MIRFOKRMAIFFLTISILIICGDIEKIISFGIEFDDILCNLFFI
GGLIICGFIYEWAKRQIKYYFELLKEDENDALEYIIVDIERVRGVVYWIAFYIIGVKC
KLRRPOYKCKDNDNYEYLEFIYKRNGEFKDFIAYADKKNDYKKSIESTNEITECEEK
AEKKRKRKRLNEYVNSRMELAKKYEEEAKNKM"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /product-"TPR-repeat-containing protein"

protein_1d-"AAK78901.1"

/db_xref-"GI:15023804

/translation-"MENENVLRAQELNNRGLIFIERNAGDEALKYFNKAIKEDENFKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /product="Predicted membrane protein; CF_1 family"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IVVPSFFVWMGVDGTEEWHKANYNLKEDALMVAANYFSDLALKVLSKK"
                                                                                                                                                                                                                                                                                                                                                                                        /product="Metal-dependent amidohydrolase"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /product="Predicted membrane protein"
/protein_id="AAK78900.1"
/db_xref="G1:15023825"
                                                                                                                                                                                                                                                                                                                                                                                                              /protein_id="AAK78897.1"
/db_xref="GI:15023822"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /protein_id="AAK78899.1"
/db_xref="GI:15023824"
'protein_id-"AAK78896.1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4594. .4863
/gene="CAC0922"
/codon_start=1
/transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /gene="CAC0922"
/4594, .4863
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /yene="CAC0923"
/gene="CAC0923"
5056. 6340
                                                                                                                                                                      3104. .4252
/gene="CAC0921"
3104. .4252
                                                                                                                                                                                                                                                                                                                    /codon_start=1
/transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /codon_start=1
/transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /codon_start=1
/transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /transl_table=11
                                                                                                                                                                                                                                       3104. .4252
/gene="CAC0921"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                056. .6348
gene="CAC0923"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6345. .6923
/gene="CAC0924"
6345. .6923
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5345. .6923
/gene="CAC0924"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7003. .8043
/gene="CAC0925"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7003. .8043
/gene="CAC0925"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /codon_start=1
                                                                                                                                         EVSNWAWNGSLF"
```

gene

CDS

gene

CDS

3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	483 513 543 573 603	633 663 693 717	753 777 810 837	870 897 930	990 1014 1050 1065	1110 1125 1170 1185	23 24 30 34	1365
AAAACAGAAAAAGATTTAATATATGCTCGTACAGATATGGGAGGAGCATATAGATGGGAC gccgccaacgggcggtggatccctctctggattgggtggatgga	GGTTTGTACACAAATGATTGGCAAGATGAAAATGCATATATTCTTAGTTCACAGGATAAA ggcgcaacgtggcaaatacgcccctgccgttcaagcttggcggcaacatgccgggcgt	ccgagcgacaagggctctggagaagcacagattccggcgacctggtcccagatgacg	gatattcaaggcgtcgtctgggtcgctttcgacaagtcttcgtcatcgctcgggcaagcg [11]	acggcgcgcgcgcggcggcgggtgccggccggcttcatcccgcacaag	tatgacgggagctccggcgacgtctggaaattctcggtgacctccgggacatggacgcga		agttatcccaatcgaagcttgcgatatgtgcttgacatttcggcggagccttggctgacc	TAACTCGGATAGAATGTTCTACGGTACAGGTGCAACTCTT
	GGTTTGTACACAAATGATTG ggcgcaacgtggcaataac	ccgagcggcaaagggctctg AGAAGTGGTAATGGACTTTG 		p-p p-p	η — Ε α — α			
250 274 310 334 364 394	424 454 484 514 544	574 604 634 664	. 694 718 754 778	811 838 871 898	931 955 991 1015	1051 1066 1111 1126	1171 1186 1231 1246 1288	1306
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	90 OX 90 OX	Oy Db Oy Db	0y 0y 0b	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	Oy Oy Ob	Oy Oy Ob	Oy Oy Oy	q

1842 2082 2142 2202 2262 2322 1467 1485 1527 1545 1587 1605 1647 1764 1785 1824 1884 1902 2004 2064 2124 2184 2244 1963 AGTTCACAAACAGGTCTTCCTACAAAAGGCAAAGGAATCTTTAAGACTGTTATAGGTCAT 2022 2245 tacggaaattggggacaagcaatcaccggtgaccacgcgaatttacggcgggtgtacata 2304 aatccgtcgatcatcgttcgcgctggaagtttcgatccatcgagccaaccgaacgacagg 1528 tcgacgatcttcacgtcaccggtgttcacgaccggcaccagcgtcgactatgcggaattg tegeaaagtgtteeegeeaatgeeeagateegeteagaeegggtgaateeaaagaettte 1945 gtcgcggccggtcttccgagcagcggtgccgtcggtgtcatgttccacgcggtgcctgga aaagaaggegatetgtggetegetegagegggetttaceaeteaaeeaatggegge 2023 GAAGGTGATATTTGGATTGCTGGAGAAAAGACGGATTATGGCATTCAACTGATTCTGGT 2065 agcagttggtctgcaatcaccggcgtatcctccgcggtgaacgtgggatttggtaagtct 2143 AAAACAGATGATGGATATCCAGCTATATACATGGATGCTACAATTGATGGAACAGCGGGA 2185 gegtacegeteegacgactgtgggacgacetgggtactgateaatgatgaceageaceaa 2203 ATATTCCGTTCAGATGATGAAGGTGCAACTTGGGTAAGAATAAATGATGCTCACCAA gtcaaaggattggaggagacggcggtaaacgatctcatcagcccgccgtctggcgccccg 1426 GCAAATGGAATTGAAGAATGTGCAGTAAATGATGTTGTAGTACCAACAAAAGGCGCACAA 1546 AGTAAAATGATGACAACACCTAACTTTAGTGCTACAACAAGCATTGATTATGCAGAAAGT ggggtgacgacggcaccgtcgccgcatcggccgacggctctcgtttcgtctggggct 1726 GGAGTITIATAAAGCAGGAACTGTAGCAGCTGGTGCTGATGCTAAAACTATAGTATGGTCT cccggcgatcccggtcagcctgtggtacgcagtcggatttggcaactcctgggctgct 1843 IGTICIGGICTICCACAGGGAGCAAAGGITAGAICAGACCGIGTAAAICCAAAGAAATIC gcgcccgggtcgtcatacccagccgtctttgtcgtcggcacgatcggaggcgttacgggg 1468 ctcatcagcgctctcggagacctcggcggcttcacccacgccgacgttactgccgtgcca 2263 TAIGGITCCCCAGACTACTGCATAACTGGAGATCCTAACAAATATGGACGAGTTTTTGTT 1408 1486 1765 1825 2125 1705 2005 QQ g g qq qq Q q Qy Db Q QY Db Qy Db Qγ qq QY Db δλ g δ qq Óγ g δý δ ò g ò δ δ οy δ

PLN 08-AUG-2000

linear

mRNA

2579 bp

ABI292929

RESULT 5 ABI292929 LOCUS

```
1083 CGGGGGTGCATCTTGGACTCGCTGTGGGCAGTCTACCCGACTCTCAACAATA 1142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1320 GGACAGTGCACACAAACGTCACCATCAAGTCGCTCGCAGATGGTGTCGAGGAAACCTCCGT 1379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        954 ctggaaatteteggtgaeetecgggaeatggaegegaateageeeggtaeettegaegga 1013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1143 CTACAGTTATAACGCTGCCCTGGCGCCGTGGATAGGACCCAACATCGTTGATGTCACCCC 1202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        792 cercecadercaaccaacraderrececcacaraaggeeerrerarececrrergagag 851
   TCTACTTGATTTCGCTGATGATTCTCGGTGGAATTATTGGGGCGTCGATGCTCTTGCAAC 311
                                                                                                                                         477 cctgccgttcaagcttggcggcaacatgcccgggcgtggaatgggcgagcgggcttgcggt
432 TITCCATTCAAGCTGGGTGGTAATAGCCTGGCGGGGGGGGGAATGGGAAGGGAGCGTGTGGTG
                                                                                                                                                                                                                                                                                            T------CTCTTCTTTGGTTTCGGAGGTCTAGCTTAGATACACAGAAAAACGGGAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               gataatggtggcaacccagatatcgtggtggccggacaccataatctttcggagcaccga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1023 TGTCATGGTTGCAGCCCTCAACTCGTGGTGGCCTGATGGTCAGATATTCCGCTCGACCAA
                                           agacccgatcaatactaacaaggtatgggccgccgtcggaatgtacaccaacagctggga
                                                                             TGATCCCGTACAACCCAACCGTCTTTACCTAGCGACTGGGTTGTATACCAACTCGTGGGA
                                                                                                                     cccaaacgacgagcgattctccgctcgtctgatcagggcgcaacgtggcaaataacgcc
                                                                                                                                                                                                                                                                      ggatccaaacaatgacaacattctgtatttcggcgccccgagcggcaaagggctctggag
                                                                                                                                                                                                                                                                                                                                              597 aagcacagattccggcgcgacctggtcccagatgacgaactttccggacgtaggcacgta
                                                                                                                                                                                                                                                                                                                                                                                                                           cattgcaaatcccactgacacgaccggctatcagagcgatattcaaggcgtcgtctgggt
                                                                                                                                                                                                                                                                                                                                                                                                                                                       CCTTCCCGATCCATCAGATAGCACCGGCCTGAATTCTGATAAGATTGGGATCGCATGGGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     cgctttcgacaagtcttcgtcatcgctcgggcaagcgagtaagaccattttgtggggcgt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             777 ggcg---gatcccaataatccggtcttctggagcagagacggcggcggcgacgtggcaggc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ggtgccgggtgcgaccggcttcatcccgcacaagggcgtctttgacccggtcaacca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1014 cacggccaacgactactttggttacagcggcctcactatcgaccgccagcacccgaacac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1314 tgatcggatgctctacggaacaggcgcgacgttgtacgcaacaaatgatctcacgaagtg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1260 GAATCATTGGCTCTACGGCACTGGAGAAACTATCTATGGAAGTCGTGACCTCCTCAAATG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ggactccggcggccagattcatatcgcgccgatggtcaaaggattggaggagacggcggt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 732 TGCGAACAAGGGCTCAAACAGCATTTATATATCTACGAATGGCGGTTCCTCTTGGAGTGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              894 cgtgctctatattgccaccagcaatacgggtggtccgtatgacgggagctccggcgacgt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              912 TTACAAGTACAGCATTGACACTGGCGCATGGACTGACATAACACCCGTCAGCGGAAGCGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  atatgtgcttgacatttcggcggagccttggctgaccttcggcgtacagccgaatcctcc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1074
                                       357
                                                                                                                   417
                                                                                                                                                                                             477
                                                                                                                                                                                                                                                                                                                                                                                                                          657
                                                                                                                                                                                                                                                                                                                                                                                                                                                             612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      834
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       972
                                     Q
                                                                         g
                                                                                                                 ò
                                                                                                                                                     q
                                                                                                                                                                                           δ
                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                              \delta
                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ob
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ΩD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /trānslation="MPSRVKFWITLTLGTFSLVYGQAASQSYRWQNVKIGGGGGFVPG
IVPRPSEKGLAXARTDIGGAYKLNADDTWIPLLDFADDSRWWYWGVDALATDPVQPNR
IVTANGLYNNSWDPSNGHILISDSYGKTFDAAPLPFRLGGNNPGRGWGFRLVDPNLN
SILYFGARSGNGLWKSTNSGRSWSKVTSFTDTGSFVPDPSDSTGLNSDKIGIAWVTID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KASGSSGSATPRIFVGVANKGSNSIYISTNGGSSWSAVAGQPTSFLPHKGVLSPSERA
LYITYSDGGGPYDGTSGAVYKYSIDTGAWTDITPVSGSDLFFGFGGLALDTOKNGTVM
VAALNSWWPDGQIFRSTNGGASWTPLMDWAVYPTLNKYYSYNAALAPWIGPNIVDVTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GNLQIGWMMESLSIDPFDSNHWLYGTGETIYGSRDLLKWDSAHNVTIKSLADGVEETS
VQALISPPSGPPLVSAIGDVGGFVHQSLTRAPSAQFTNPKWSTSADIDFAGNNPSNVV
RIGTGDSTSGKQVAISSDYGVTWNQHFGAPDNVQGGKVAISADADIILWRTNGNGVMV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SRNQATFNIFYGASGKTFYVSTDNGKTFSAHGSLGSATSVFDITVHPSVSGDIMASTD
KGLFHSTDSGATFSAISGITQAMGVALGAPRSTGGYPAVFAAANYGGEVAYFRSDDRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VNWVKINDAAHGFGAASANCMAADPRVYGRYYIGTNGRGIFYGDVAGSAPPTSSATS
TTSVSTTTTRSTTTTRSTTTTTSSGNGSPVSGFGQCGGQGWTGGTTCVTGFQCVAQNP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5
                                                                                                                   Homobasidiomycetes;
                                                                                                                                                                                             bisporus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              cggggggggggggtttgtcgacgggatcgtcttcaatgaaggtgcaccgggaattctgta 236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  cgtgcggacggacatcggggggatgtatcgatgggatgccgccaacgggcggtggatccc 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   tcttctggattgggtgggatggaacaattgggggtacaacggcgtcgtcagcattgcggc 356
                                                                                                                                                                                                                                                                    Direct Submission
Submitted (07-AUG-2000) Morales-Almora P., Microbiology, King'
COllege London, 150 Stamford Street, London, SE1 8WA, UNITED
KINGDOM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             in Agaricus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 2579;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                           Agaricus bisporus
Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
Agaricales; Agaricaceae; Agaricus.
1 (bases 1 to 2579)
                                                                                                                                                       1 (bases 1 to 22/2)
Morales-Almora,P. and Thurston,C.F.
Morales-analysis of the cellulolytic genes
Unpublished
(1 to 2579)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          734;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 327.6; DB 8
Pred. No. 8.3e-28;
0; Mismatches 734
   protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /function="putative cellulase"
/codon_start=1
                                                                                                                                                                                                                                                                                                                                                                          /organism="Agaricus bisporus'/strain="D649"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               'country="United Kingdom"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /product="CEL6 protein"
/protein_id="CAC02964.1"
/db_xref="G1:9755897"
 CEL6
                                                                                                                                                                                                                                                                                                                                                                                                                                       /clone="PMRT152"
/dev_stage="mycelium"
                                                                                                                                                                                                                                                                                                                                                                                                                        /db_xref="taxon:5341"
                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
 for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /gene="ce16"
/product="CEL6"
662 c 631
Agaricus bisporus mRNA
AJ292929
AJ292929.1 GI:9755896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28. .2448
/gene="cel6"
97. .2445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /gene="cel6"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /gene="cel6"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    11.4%;
52.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     . 2448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 52.6
Matches 845; Conservative
                                                                           Agaricus bisporus
                                                                                                                                                                                                                                                Morales-Almora,P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           96
                                                         cel6 gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sig_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
 DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BASE COUNT
ORIGIN
                                                                                                                                                     REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
                 ACCESSION
VERSION
                                                                                             ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   297
                                                       KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                        JOURNAL
                                                                                                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
```

```
giggigaataccgggicgicgicgiggattigicgacggigacggigcggiactggitc 2742
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2766 ATTGTGAACCGCGGACGACGAGCGTACCGTTGTCGGAGCTTACGATTCGGTACTGGTAC 2825
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2743 accegggatggtggctcgtcgacactggtgtacaactgtgactgggcggtcgatcggtgt 2802
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2563 tcaccgtcttcgtcgccgagcccaacgccgtcgtcgtcgccggtgtcgggtggggtgaag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2649 ACGGGGACACCGACACCCACGCGGACGCCGTCGGCCG---GCGGCAACCTGGTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            beta-1,4-xylanase catalytic domain"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      gtgcagtataagaataatgattcggcgccgggtgataatcagatcaagccgggtttgcag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2443 tegeogagetegagtecatectegtegeegteteegtegeegteaceategeegagteeg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /note="Region: cellulose-binding domain typeIIIb"
2520. .2696
/gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    cellulose-binding domain typeIIIb"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 131.6; DB 1; Length
Pred. No. 3.7e-06;
0; Mismatches 164; Indels
                                                                                                                                                                                                                                                                                                                                                  384. .482
/gene="xynA"
/gene="xynA"
/gene="xynA"
/note="Region: thermostabilizing domain"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /note="Rêgion: proline-threonine linker"
2076. .2519
/gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    'note="Region: proline-threonine linker
                                                                                                                                                                                                                                                                                                                                                                                                                                        483. .3146
/gene="xynA"
/product="beta-1,4-xylanase XynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    951
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /note="Région: L
1947. .2075
/gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note="Region:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           966. .1946
/gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2697. .3146
/gene="xynA'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      825 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 60.13
Matches 256; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2697.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     misc_feature
                                                                                                                                                                                                                                                                                                                                                                                         misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             misc_feature
                                                                                                                                                                                                                                                                                                                                                   sig_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                             mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QΣ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /translation="QVGCSNLRGSFVKLSTGRTGADYYIEITFTSGAGSLAPGASSGD
IQVRINKNDWTNYNEANDYSYDPTKTSFADWNRVTLYRNGQLIWGVEP"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BCT 01-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2 (bases 1 to 3237)
Sunna, A., Gibbs, M.D. and Bergquist, P.L.
Direct Submission.

Submitted (29-cor-1999) Biological Sciences, Macquarie University,
North Ryde, Sydney, New South Wales 2109, Australia
Location/Qualifiers
1 . 3237
/organism="Caldibacillus cellulovorans"
//db_xref="taxon:74586"
                                                                                                                                                                                                                                                                                                                                                                                                                                               3237 bp DNA linear BCT 01-DEC-20 Caldibacillus cellulovorans putative type IIIb cellulose-binding domain gene, partial, cds; and beta-1,4-xylanase XynA precursor AF200304
                                                                           1494 eggetteacceaegeegaegttactgeegtgeeategaegatetteaegteaeeggtgtt 1553
                                                                                                                                                                                                                                                                                 1674 caagaactggttccaaggcagcgaacctggcggggtgacgacgggggggcaccgtcgccgc 1733
                                                                                                                                                                                                                                                                                                                 1611 CGTTACTTGGAATCAGCATTTTGGAGCTCCTGATAATGTTCAAGGCGGAAAGGTTGCCAT 1670
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /evidence=not_experimental
/product="putative type IIIb cellulose-binding domain"
/protein_id="AAF61648.1"
/db_xref="GI:7385019"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 (bases 1 to 3237)
Sunda,A., Gibbs,M.D. and Bergquist,P.L.
A novel thermostable multidomain 1,4-beta-xylahase from
(Caldibacillus cellulovorans' and effect of its xylan-binding
            aaacgatctcatcagcccgccgtctggcgccccgctcatcagcgctctcggagacctcgg
                                          1380 CCAAGCCCTTATTTCACCTCCTTCCGGACCGCCTCTTGTTTCTGCTATTGGTGACGTTGG
                                                                                                1440 CGGTTTGCCACCAATCACTTACCGTGCACCCAGCGCCCAGTTTACCAATCCCAAGTG
                                                                                                                                                                               1500 GTCGACCTCGGCTGACATCGATTTCGCTGGCAATAACCCTAGTAATGTCGTTCGCATTGG
                                                                                                                                                                                                               1560 TACTGGTGATAGTACTAGT-----GGCAAACAGGTCGCCATCTCTTCCGACTATGG
                                                                                                                                               1554 cacgaccggcaccagcgtcgactatgcggaattgaatccgtcgatcatcgttcgcgctgg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bacillus/Staphylococcus group; Alicyclobacillus group; Caldibacillus.
                                                                                                                                                                                                                                                                                                                                                                       1671 CTCCGCTGATCATCATCTCTGGCGTACAAATGGTAAACGGT 1716
                                                                                                                                                                                                                                                                                                                                                     1734 atoggoogacggototogtttogtctggggctcccggcgatcccggt 1779
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bacteria; Firmicutes; Bacillus/Clostridium group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        domain on enzyme activity
Microbiology 146 (Pt 11), 2947-2955 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /gene="xynA"
/note="multidomain protein"
/codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Caldibacillus cellulovorans. Caldibacillus cellulovorans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /codon_start=1/trans1_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AF200304.1 GI:7385018
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        384. .3149
/gene="xynA"
384. .3149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /note="orf1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        20519260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11065373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE
PUBMED
          1434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                           AF200304
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
                                                                                                           g
                                                                                                                                               ò
                                                                                                                                                                           g
                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                 염
                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                            ò
          ò
```

ς,

Gaps

9

Length 3237;

2562 2648 2622 2705 2682

```
/codon_start=1
/product="1629-caps.id"
/protein_id="Aac59001.1"
/db_xref="G1:2034.904"
/db_xref="G1:2034.904"
/translation="MMERQYQSVRSYLINNQHNAIAAGPFLQRVAGPEAHSVGRNVGD
RAYRLINGAVLDLIEKLAEDIYADTAXMQADQPEASSRHFATLURMXLLLIGVQDPDAR
RNLNSVLARIEALLRVDVVNDAEVNVLSGDFYEEYSKY1SYQQTFAQTPTASASQQTQ
TSLPRQTQASLERQYQPFDQPEMVSPPSFVHTTPALILAQTPTASASQQTQ
TSLPRQTSLERQYQASPEDQYPFEDQPEMVSPPSFVHTTPALILAQTPQPPATDYFSRP
SDEFVYVPGKERAVPDTRFKPPVPFRPEHLKSRPSSVATNAAGATPVAPPPPFPSRD
VTTSMPPPPPFFSADVTTSMPPPPPNVDLATSMPPPTNNA
INNLLIDANVAETNKNAGDNRSALLQSIKQGKTLKKTQPADGARATDRFSTLLSSIRQ
GKTLKKLRKLERQEGSSTQLLKOVDTTPKTKTILKNFVTNIDRISKQEGEEKDRLDTIT
KRRPAVEHTDGNSTGNNSDDMRDD"
                             OPU75930 131995 bp DNA circular VRL 26-OCT-2000
Orgyia pseudotsugata nuclear polyhedrosis virus complete genome.
U75930
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /translation="moral"/
RKTIAAHSFGADEINVHDLASDHPSFVDDCAVLAPKVVNGRFGKMDVLHHRPTTSKLFL
RKTIAAHSFGADEINVHDLASDHPSFVDMYFCYSSPTAWAIVMDYVECPDLEFTLQTQ
GALDMALVWIVRQLCDALNDLHNATGYIHNDVKLENVLYFGARDRYYLCDYGLCKRE
HSPVHDGTLEYFSPBETRRHHYARSFDWYAVGVLAYKLLTGGRHPFBRSADBYLDLAS
MRRQQYNDPAALKNVRNLMARDFVFCLTRFNFECRSTDYKQIAKHSFLASRHDYI"

    (bases 1 to 131995)
    Ahrens, C.H., Russell, R.L., Funk, C.J., Evans, J.T., Harwood, S.H. and
Rohrmann, G.F.

                                                                                                                                                                                                                                                                                                                           The sequence of the Orgyia pseudotsugata multinucleocapsid nuclear polyhedrosis virus genome Virology 229 (2), 381-399 (1997) 97271300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1. .131995

/organism="Orgyia pseudotsugata single capsid nuclear polyhedrosis virus"

polyhedrosis virus"

/db_xref="taxon:10450"

complement(123. .947)

/note="ORF1: pk-1, AcMNPV ORF10 homolog"

/codon_start=1

/product="protein kinase 1"

/protein_id="AAC59000.1"

/db_xref="Gr:1911247"

/db_xref="Gr:1911247"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Direct Submission
Submitted (26-0cr-2000) Oregon State University, Agricultural
Chemistry, Corvallis, OR 97331-7301, USA
Sequence update by submitter
On Oct 26, 2000 this sequence version replaced gi:2934903.
similar to Autographa californica nuclear polyhedrosis virus
(AcMNPV) complete genome: GenBank Accession Number L22858.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Submitted (23-007-1996) Oregon State University, Agricultural Chemistry, Corvallis, OR 97331-7301, USA (bases 1 to 131995) Rohrmann, G.F.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            University, Agricultural USA
                                                                                                                                                              Orgyia pseudotsugata single capsid nuclear polyhedrosis orgyia pseudotsugata single capsid nuclear polyhedrosis Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /note="ORF3; ph, AcMNPV ORF8 homolog"
/codon_start=1
/product="polyhedrin"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /note="ORF2; AcMNPV ORF9 homolog
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Direct Submission
Submitted (06-MAR-1998) Oregon State
Chemistry, Corvallis, OR 97331-7301,
Sequence update by submitter
4 (bases 1 to 131995)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    complement (2533.
                                                                                                                                                                                                                                                                                                                                                                                                                     2 (bases 1 to 131995)
Rohrmann, G.F.
Direct Submission
                                                                                                                  U75930.2 GI:11024985
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      source
                                              LOCUS
                        OPU75930/c
                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                         JOURNAL
MEDLINE
                                                                                            ACCESSION
                                                                                                                                                                                                                                                             REFERENCE
                                                                                                                                                                                                                                                                                      AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
                                                                                                                VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REMARK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REMARK
                                                                                                                                                                                                                                                                                                                               TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Analysis of the region from the SstI I fragment that was crosshybridized with the repeated DNA downstream of the polyheadrin gene revealed a sequence made up predominantly of repeats of the sequence GGCG or the closely related variant GGTG, and a dinucleotide which results in amplification of groups of six
                                                                                                                                                                                                                                                                                                                               VRL 03-FEB-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 repetitive GC-rich region.
Multicapsid nuclear polyhedrosis virus of Orgyia pseudotsugata DNA,
clone cosmid 58.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Chen, D.D., Nesson, M.H., Rohrmann, G.F. and Beaudreau, G.S. The genome of the multicapsid baculovirus of Orgyia pseudotsugata: restriction map and analysis of two sets of GC-rich repeated
2826 ACGGTGGACGGCGACAAGCCG---CAGGTGTTCAACTGTGACTGGGCGTGGGTCGGATGT 2882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2348 cgccgtccggatcgccgtctccgtcggtgagtccgtcggcttcgccgagcctgagcccga 2407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2408 gecegagecegagetegecategeegtegeegtegeegagetegagtecatectegt 2467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     cgccgtctccgtcgccgtcaccatcgccgagtccgtctcggtctccgtcaccatcggcgt 2527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /organism="Orgyia pseudotsugata single capsid nuclear polyhedrosis virus"
/db_xref="taxon:10450"
in 163 c 289 g 186 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Orgyla pseudotsugata single capsid nuclear polyhedrosis virus
Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
                                                                                               298 GGCCATCACCAACGCCATCACCAACGCCGACGCCATCGCCGACCCCAACCCGA
                                                                    2803 gggaatatccgcgcctcgttcggctcggtgaacccggcgacgccgacggcgacacctac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NPORPT2 11near VRL 03 Multicapsid nuclear polyhedrosis virus of Orgyia pseudotsugata(OpMNPV), GC-rich repeated DNA from the Sstl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 14; Length 749;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3.6%; Score 103.6; DB 14; Length 62.2%; Pred. No. 0.0066; Live 0; Mismatches 99; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Virol. 69 (Pt 6), 1375-1381 (1988)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               289 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       up upstream of PstI site.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              egcegtegtegtegeeggtgte 2609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        178 GCCATGGCCGACGCGGCG 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Nucleopolyhedrovirus.
1 (bases 1 to 749)
                                                                                                                                                                                                                                                                                                                                                                                                                                           GI:222216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                    fragment.
D13755 D00297
D13755.1 GI:2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                nucleotides.
                                                                                                                                                                                                                                                                                                                                                   Multicapsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                          CTTGAG 2948
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sednences
                                                                                                                                                              2863 ctgcag 2868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Sim:
Matches 163;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                           LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BASE COUNT
ORIGIN
                                                                                                                                                                                                          2943
                                                                                                                                                                                                                                                                                 RESULT 7
NPORPT2/c
                                                                                                                                                                                                                                                                                                                                                                                                                     ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JOURNAL
MEDLINE
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURES
                     g
                                                                                                                                                         ò
                                                                                                                                                                                                        g
                                                                 ò
                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
/traislation="MAPCKYTPERVQMMMDAIAYNDSRRLAFMTDRPRWVHAHNFPDS
AAQLEAXIVRNSISDYHVRPLEEGGREWYIDADFRDCADKAELMLKYWGATAFMLFF
EGREDAVQT IMPSGNRFHMLKFCGKFKMDAPKSLREHWFNVFKQPAKLVSGDIRPG
SFADCVYRRAYHMXIGDAREDLYLRYWPDVDRUYFCNARKQIRAPFSYNYKGGDYSRCL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /protein_id="AAC59014.1"
/db_xref="GI:1911261"
/translation="METAQPPISYAPPKRGAVCAYVRTVYTTTVSDSGGNNEDRLTQ
                                                                                                                                       /translation="MGLSSKLLVYAYYGSYNLPHDRYGESYHLYRIVHEHLTNTYVSN
ASCVRRDIATARCLNSGHLCFDVARQLLDVSEVAARLSAWFRCGDATGLCADMQRALA
DIDRHAPLARRYGRRANIFALDAIADIPSDVTNNLQGIIGRFWHFPRCSGLARVADVF
                                                                                                                                                                                                                           DPDIRADGWWYHKFCVLTYMHLVACGAVPAGSAFRLRDAVAKHIGPNDEGNCAPAIAA
VYGRPCAIGREHFAHHKTACMHILFQFMRNDLTPADERHPCFGVIKDFGRQCKDTYTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /translation="MFSWMFGWWNAADEQVNAEFDEQAYRRYAVDQRAHSDLVRRDVFRCHPFVFKFRYVLDETAGRCCSVVDFCKGLKISHDLLQRCNFDRQHVRQLNELVLGAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PAQPDSLGSLFATKHGLVQLLQQFSFANKNEVLLAVGANKDHDRDNLLDKIEAVLNHV
KTLNTNSDKFISAHKSFKLEVGARFEQFEQRLQTLDTKLNALQCAAPTRTAPGVVFPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /traislation="MyFLITALFLLATGARAASILAVLPTPAYSHHVVYRAYVHALVK
NCHNVTVIKPQLLDYAVQDECGRVEQIDADMSAQQYKKLVASSGVFRKRGVVADETTV
TADNYMGLIEMFKDQFDNANVRRFLSTNRTFDAVVVEAFADYALVFGHLFRPAPVIQI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APGYGLAENFERRRÂVARHPLHYPTFGAAALTRRGGALSEWRLLNEFELLARRSDELL
KQQFGKSTPTIRQLRDNVQLLLLNLHPVYDNNRPVPPSVQYLGGGLHLAQALPQRLDA
PLERRLNESVDGAVYVSFGSGIDTNSIHAEFLQMLLDTFANLNNYTVLWKVDDAVAAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VALPRNVLAQKWFSQTAVLRHKNVVĀFVTQAGĪQSSDEĀLQĀRVPMVCLPMMGDOFHH
VREKLQOFGYARALDTRAVSĀPOLQLĀIREVIABGEĀYRARIDKLRĀVVEHDĀĀPDEKĀ
VKFTERVIKFNIDVIMPĀRSLKTTĀANMĀYSDYFVRFPL"
11547. . 12140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DVTKHPHLAVFMGRVEDRGVTQIAFARGQEEHFRKRKLEFEEGMDVDVRGRAPNPLLA
.VHCIKEEFANGGHKIRRLPKKVIEVDCAVNVAKDIVKKAILNKT"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                 LRTHADALYIHGTTDRQKNALFDLLCCVNASDIDADCYDCVVNKFYATQNKKYKM"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      75588 CGCCATCGCCGACGCCATCACCAACGCCGACGCCATCGCCGACGCCCATCACCAACGCCGA 75529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             75469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2348 egcegtecegategeegteteegteggtgagteegteggettegeegageetgageeega 2407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2408 gecegageeegageagetegeeategeegtegeegtegeegagetegagtecatectegt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 131995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /product="ecdysteroid UDP-glucosyltransferase"
/protein_id="AAC59013.1"
/db_xref="G1:1911260"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     complement(9094. .9825)
/note="ORF13; lef-1, AcMNPV ORF14 homolog"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 103.6; DB 14; Length
Pred. No. 0.0019;
0; Mismatches 99; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOQLQORIKACSAGCLAGGTPFTSK"
9893. 11362
/note="ORF14; egt, AcMNPV ORF15 homolog"
/codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /codon_start=1
/prodouct="late expression factor 1"
/protein_id="AAC59012.1"
/db_xref="G1:1911259"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note="ORF15; AcMNPV ORF16 homolog"
                                                                                                                                                                                                                                                                                                                  complement(8180. .9142)
/note="ORF12; AcMNPV ORF13 homolog"
/codon_start=1
/note="ORF11; AcMNPV ORF11 homolog"
                                                                                                                                                                                                                                                                                                                                                                                                                              /protein_id="AAC59011.1"
/db_xref="G1:1911258"
                                                                              /protein_id="AAC59010.1"
/db_xref="G1:1911257"
                                                            /product="unknown"
                                                                                                                                                                                                                                                                                                                                                                                                      /product-"unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /product-"unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Sim
Matches 163;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                     CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                   /translation="MPDYSYRPTIGRTYVYDNKYYKNLGSVIKNAKRKKHLLEHEEDE
KHLDPLDHYWVAEDPFLGPGKNOKLTLFKEIRNVKPDTWKLIVNWSGKEFLRETWTRF
VEDSFPIVNNOGEWDVFLVVNMRPTRRNKCYKFLAGIALRWDCDYVPHEVIRIVEPSY
VGMNNEYRISLAKKGGGFIMNIHAEYTNSFESFVNRVIWENFYKPIVYIGTDSSEEE
ELLIEVSLVPKVKERADAAPLFTGFAY"
complement(3350. .3541)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /translation="WEGGARVESIRLTQEFKENVAHVDHLMGLRALIDGKVTSADV
RRFGFLSRNVLVSACMAVNVQVYAPDATIDMRHQPTIYFRVCQNCHAMADVPAPDDHS
IARYLLAEGGAVLVIDHPLDYFGETEEGVNELLEVQRINAGGDL"
complement (5669. 6151)
/note="ORF9; ptp-2"
/codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /translation="MNRPTMRNTAAVTTDYDREQLRRELNSLRRSVHELCTRSATGFD
CNRFLEAGDRAPAVIVKAAANGGQHSSLICDKV"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /translation="MLTSRRWAVIYTRTSERNCGGWCTNGVRRROVHLPSVRARAM
PEDRWHEYTANGGAVIEGTRLLCFKVPENABLEFGYTSDEDRWTASVLARHSALGAVI
DEDRWTARYDGAVKMGLLYFRYPGRAVDDDIVAEFIETWPDEFRRCANLOW
HWTHGLNRSGYLVCRYWVERLGVSPTDAIARFETARGHKIERTNYLQDLLARKHVRGO
                                                                                                                                                                                                                                                                                                                                                                     /translation="MRKNNGASOYPEHLTTVHKRHRLQLSCVWCDDQRSWDPHSAKGL.
MFEHMFCCLRIYVQQXIKH"
complement(3698 . 4093)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /translation="mgasquseqkIlkLFyrwssqtgaalddekdLHCLYDLERFvGa
HLNKRADDKaRKKKCAEKAALKRVEIAADRHMLEAAAAPACADDGRWSTLSRAQLDDI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /trānslation="mervwnPaagiDGLKRSETYLVDPHDFVGVLTLSPYTVFERGLF
VRMSGMRLLALLAAPKPQEPQPAVRRFPQRSRRNVCLKACADGAQSLAKVLAARVSMP
PCMSKTMADLSSAPRGNMYRKRFEFNCYLANVITCTKCKTACLIGALLHFYRMDAKCV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /translation="MYEASQIDEHVFVGGYYGDNEAMLRFIENHAIASVISLIDSDVA
PIRHALGLPVGDHIHVYCEAAPTCAALPNAMPALYDYMVRRIGEGKRVLVHCYAGASR
SAALAVYYLMRSRQMPYQDALNAVQSKRRVAINDHFVRFLATRCSYRFVNDELKPQIV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GEVTHLLIKAQDVYKPSNCAKMKKVTKLCPQASMCKGLNPICNF"
complement(4742. .4975)
/note="ORF7; AcMNPV ORF5 homolog"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /product="protein tyrosine phosphatase 1"
/protein_id="AAC59009.1"
/db_xref="GI:1911256"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /product-"protein tyrosine phosphatase 2"/protein_id="AAC59008.1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             homolog"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                complement(4125. .4739)
/note="ORF6; lef-2, AcMNPV ORF6 homolog"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /product="late expression factor 2"
/protein_id="AAC59005.1"
/db_xref="G1:1911252"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5020. .5460
/note="ORF8; AcMNPV ORF4 homolog"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note="ORFi0; ptp-1, AcMNPV ORF1 /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AREKEIVDRIHRLQLKQDSLLKQDRLKKR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /protein_id="AAC59006.1"
/db_xref="G1:1911253"
                                                                                                                                                                                                                                                                                                         /protein_id="AAC59003.1"
/db_xref="G1:1911250"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /protein_id="AAC59004.1"
/db_xref="GI:1911251"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /protein_id="AAC59007.1"
/db_xref="G1:1911254"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .6791)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /db_xref="GI:1911255"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /product-"unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /product="unknown"
                                                                                                                                                                                                                                                                                          /product="unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /product-"unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     complement (6129.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /codon_start=1
                                                                                                                                                                                                                                                            /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /codon_start=1
                                                                                                                                                                                                                                 /note="ORF4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6819. .7814
                                                                                                                                                                                                   CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                    CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CDS
```

us-09-917-376-2.rge

```
/translation="TPTPTPTPSAAPTPPSAGGSLVVQYRAADTNAGDNQLKPHFRI
VNRGTSSVPLSELTIRYWYTVDGDKPQVFNCDWAQVGCSNVRGSFVKLSTGRTGADYY
IEITFTSGAGSLAAGASSGDIQVRINKNDWTNYNEANDYSYDPTKTSFADWNRVTLYR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /protein_id="AaF22274.1"
/db_xref="G1:6651327"
/tb_xref="G1:6651327"
/translation="MNRIALBALSGMLAMVLIAAMLAYVPKPEPAEAHGGMVFPATRT
YACYVDGKYHGNGGDLMAINPACLDALAISGNYQFWWWFGNIISNAGGRHREIIPDGK
LCGPTASFDGMNQARTDWWTTRLQPGAIITVRVNAWAPHPGTWYLXVTRDGWDPTQPL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PLVIGEFGHMHSDGDPNEQAIVQYAKQYNIGLFGWSWSGNGGGVEYLDMYTNFNANSP
TAMGTWFRTNAIGTSTSPTPTPTPTPTPTPTPTPSAGGNLVVQYRAADTNATDNQL
KPHFRIVNRGTSSVPLSELTIRYWYTVDGDKPQVFNCDWAQVGCSNLRGSFVKLSTGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TGADYY IEITFTSGAGSLAPGASSGDIQVRINKNDWTNYNEANDYSY DPTKTSFADWN
RVTLYRNGQLVWGVEP"
                                                                                                        Macquarie University,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /product="multidomain beta-1,4-mannanase precursor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61. .510
/note="Region: cellulose-binding domain type IIIb"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               domain type IIIb"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /product="multidomain beta-1,4-mannanase"
1349. .1480
/gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note="Region: proline-threonine linker" /evidence=not_experimental
                                                                                                                                                                                                                                                                                                                                                                                                                     1. .60
/note="Region: proline-threonine linker"
                                                                                                                                                   1. 4567
/organism="Caldibacillus cellulovorans"
/db_xref="taxon:74586"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note="Region: proline-threonine linker"
Environ. Microbiol. 66 (2), 664-670 (2000)
                                                              Sunia, A., Gibbs, M.D. ....
Direct Submission
Submitted (29-JUN-1999) Biological Sciences,
North Ryde, New South Wales 2109, Australia
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /note="Region: cellulose-binding
                                                   2 (bases 1 to 4567)
Sunna, A., Gibbs, M.D. and Bergquist, P.L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /evidence=not_experimental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /evidence=not_experimental
1928, .2032
                                                                                                                                                                                                                                                                                                                                                                                                                                                         /evidence=not_experimental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /evidence=not_experimental 620, .3412.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /evidence-not_experimental
                                                                                                                                                                                                                                                                                                         /protein_id="AAF22273.1"
/db_xref="GI:6651326"
                                                                                                                                                                                                                                                                     /transl_table=11
/product="unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /codon_start=1
/transl_table=11
                                                                                                                                                                                                                                                       /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /gene="manA"
/200. 710
                                                                                                                                                                                                                                   /note="orfl"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                719. .3409
/gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                  NGQLIWGVEP'
 Appl. Env
20120520
                                                                                                                                                                                                                                                                                                                                                                                                                      misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sig_peptide
                                                                                                                                                                  source
                                                                                       TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gene
                                                                       AUTHORS
                                                     REFERENCE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CDS
                                                                                                                                             FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AF163837 4567 bp DNA linear BCT 08-FEB-2000 Caldibacillus cellulovorans multidomain beta-1,4-mannanase precursor (manA) gene, complete cds; and unknown genes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 (bases 1 to 4567)
Sunna,A., Gibbs,M.D., Chin,C.W.J., Nelson,P.J. and Bergquist,P.L.
A gene encoding a novel multidomain beta-1,4-mannanase from
Caldibacillus cellulovorans and action of the recombinant enzyme on
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ·;
                                                                                                                                                                                                 PAT 16-MAY-2001
                  egecgagecegtettegteacegagecegtettegteacegtettegtegeegageeeaa 2587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2407 agcccgagcccgagcagctcgccatcgccgtcgccgtcgccgagctcgagtccatcctcg 2466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2467 tegeegteteegtegeegteaceategeegagteegteteggteteegteaceateggeg 2526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2527 tegeegagecegtettegteacegagecegtettegteacegtettegtegeegageeea 2586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2347 gegeegteeggategeegteteegteggtgagteegteggettegeegageetgageeeg
                                                                                                                                                                                                                                                                                                                                         1 (bases 1 to 390)
Gold,L., Tuerk,C., Pribnow,D. and Smith,J.Drew.
Systematic polypeptide evolution by reverse translation
Patent: US 6194550-A 7 27-FEB-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     caldibacillus cellulovorans.
Caldibacillus cellulovorans
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/staphylococcus group; Alicyclobacillus group;
Caldibacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 390;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0; Mismatches 102;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 100.8; DB Pred. No. 0.016;
                                                                                                                                                                                                 AR135147 390 bp
Sequence 7 from patent US 6194550.
AR135147
                                                                                                                                                                                                                                                                                                                                                                                                                                                     /organism="unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                75348 CGCCATCGCCGACGCCGCC 75327
                                                                      2588 egeegtegtegtegeeggtgte 2609
                                                                                                                                                                                                                                                    AR135147.1 GI:14124052
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AF163837.1 GI:6651325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              tch 3.5%;
al Similarity 61.4%;
162; Conservative
                                                                                                                                                                                                                                                                                                                            Unclassified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        kraft pulp
                                                                                                                                                                                                                                                                                        Unknown.
                                                                                                                                                                                                                                                                                                         Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local 3
                                                                                                                                                                                                                DEFINITION
ACCESSION
                                                                                                                                                               RESULT 9
AR135147/C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                           ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
AUTHORS
TITLE
2528
                                                                                                                                                                                                                                                                                                                                                              AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                             REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                  JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACCESSION
                                                                                                                                                                                                                                                                     KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AF163837
                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                       VERSION
                                                                                                                                                                                                  LOCUS
                                q
                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
127716
            Direct Submission
Submitted (31-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
Submittedshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 4, 2002 this sequence version replaced gi:18476868.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        127656
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Db 127595 GCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCTCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         127596
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2347 gegecgicegaicgecgictecgicggigagicegicggeticgecgagectgageceg 2406
                                                                                                                                                                                                                                                                             Assembly program: XGR4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 183201 bases at least Q40
Consensus quality: 183287 bases at least Q20
Consensus quality: 183384 bases at least Q20
Insert size: 183465; sum-of-contigs
Insert size: 176789; 3.5% error; agarose-fp
Quality coverage: 13.90x in Q20 bases; sum-of-contigs Quality
coverage: 15.10x in Q20 bases; agarose-fp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Db 127655 GCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCTCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2407 agcccgagcccgagcagctcgccatcgccgtcgccgtcgccgagctcgagtccatcctcg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Db 127775 genecedentecedentecedentecedentecedentecedentecedentecedenteced
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NOTE: This is a 'working draft' sequence. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   tegecegagecegtettegteacegagecegtettegteacegtettegtegecegagecea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2587 acgocotcotcotcocogtotcogotogogtogaaggtgcagtataagaataatgattco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 183665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 91.4; DB 2; Length 1
Pred. No. 0.04;
0; Mismatches 136; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note="assembly_fragment:05886"
3251. .60564
/note="assembly_fragment:00037
fragment_chain:1"
                                                                                                                                        Center: Wellcome Trust Sanger Institute
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note="assembly_fragment:06038
fragment_chain:1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              60618 t
                                                                                                                                                                        1. .183665
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
                                                                                                                                                                                                                                                                ----- Summary Statistics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /clone_lib="RPCI-23"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /clone="RP23-396N4"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       vector_side:right"
a 35803 c 35999 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    clone_end:SP6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3.2%;
56.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 56.0
Matches 173; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              51045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BASE COUNT
ORIGIN
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                    COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                 /translation="mgasgyfsrtagngyglkarcokrrcygrnllghrlfyydddef
RFFAGARKRLGRFSRRONRLGEAQNFESGTVTFYOPDDKRRIRRENIGQLFYRFQASI
GPKRLIEIAEORQARARKQASDQPQLKNAVILHFIDQHMPYGYERMDPFHKFIEKQKS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AL645600 183665 bp DNA linear HTG 01-FEB-2002 Mus musculus chromosome 11 clone RP23-396N4, *** SEQUENCING IN PROGRESS ***, in unordered pieces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2790
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2554 cogicitogicacogicitogicogagocoaacgocogicogicogicogigicoggi 2613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2730
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2494 ccgagtccgtctccgtcaccatcggcgtcgccgagcccgtcttcgtcaccgagc 2553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3134 CAGGTGGGTTGCTCGAACTTGCGGGGCAGCTTCGTGAAGCTTTCGACGGGCCGGACGGGG 3193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2837 CCGACGCCTGGGGGACATGGTTCCGGACGAACGCGATCGGGACCTCGACTTCGCCGACA 2896
                                    'note="Region: beta-1,4-mannanase catalytic domain"
                                                                                                               /gome="Region: proline-threonine linker"
/evidence=not_experimental
2960. .3409
/gene="manA"
/note="Region: cellulose-binding domain type IIIb"
/kevidence=not_experimental
/note="orf3"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2957 GGCAACTIGGTCGTACAGTACCGCGCGGCGGACGAACGCGACGACCAGCTAGAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ccgggtttgcaggtggtgaataccgggtcgtcgtcggtggatttgtcgacggtgacggtg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ccecariticecaritergaaccecegeacerceaecereccecrerceaecreaecarit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                cggtactggttcacccgggatggtggctcgtcgacactggtgtacaactgtgactgggcg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         gcgatcgggtgtgtgggaatatccgcgcctcgttcggctcggtgaacccggcgacgccgacg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9999---tgaaggtgcagtataagaataatgattcggcgccgggtgataatcagatcaag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AL645600
AL645600.8 GI:18491381
HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 4567;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 96.4; DB 1;
Pred. No. 0.027;
0; Mismatches 156;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     812 t
                                                                                                                                                                                                                                                                                                                          /product="unknown"
/protein_id="AAF22275.1"
/db_xref="GI:6651328"
                                                     /evidence-experimental
2885. .2959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1276 c 1439 g
                                                                                                                                                                                                                                                                                 /codon_start=1
/transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3194 GCGGACTACTACATCGAG 3211
                                                                               2885. .2959
/gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                  RSILTRYY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              geggacacetacetgeag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mus musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
Matches 216; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       house mouse.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1040
misc_feature
                                                                               misc_feature
                                                                                                                                                            misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 11
AL645600/c
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2791
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2731
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2851
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

```
/translation="MRKLVFSDTERFYRELKTALAQGEEVEVITDYERYSDLPEDLKT
IFELHKNKSGMWVNVATGAFIPYSFAATTINYSALYIFGGAGAGAIFGVIVGGPVGAA
VGGGIGAIVGTVAVATLNGKHHVDIEINANGKLRFKISPSIK"
                                                                                                                                                                                                                                          /codon_start=1.
/codon_start=1.
/protein_id="BAB77489.1"
/db_xref="G1:1134934"
/translation="MVAQFSSSTAIAGSDSFDIRNFIDOLEPTRVKNKYICPVCGGHN
/LINDNGKYSCYNECLHRDIREALREWIGVELESPRELDFTRAKRATYPK
VLDVDPSQIAICLIGESETTPPQPYTPPCFWSVAIRLSDSGATSQELKEIKETEYDVG
NGRRAHRFSCPCAAAPKGRVATFSVSRIDPITNKVAMKKEGFWPAYRQSEAIAIIKAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                              DGIPVLIAHEGERCVEASRLEHLASITWYGSSSBRDILHSLYOLOHNYSTROWNYDD BOTTVLAHEGERCVEASRLEHLASITWYGSSSBRDILHSLYOLOHNYSTROWNYDD BODSTGWNKQORIKEICQAGVSFVAIDLLKIQDRGDIADIFESGMTGDELVELLAYGOSTTTQKQQQSAQDTLDYADLGDEEPDSTFLQKGPTATYGERKITAGAGKLYY
WHGKGYRYSPPAVERRITYGNSYPYIKKNGDSYYTYPYARNYKREILGEWYKLRY
BIDPQLINPPGYNYRAINYGNSYPYIKKNGDSYYTYPYARNYKREILGEWYKLRY
CDRLLECLDSPQQQILLRNLGASLDLABVRRKGREPKYLLAGGGLANGKDSIRQVVS
TIYGHQGAGTSCSLADFYAYDGGRFAAALVNRRNASENPQTTRLLNGYGSLKLFAT
GNYLHSERGKGDIIEPNAAIGIFNAUDTPSWYGTIQAILDRIAALIFRKTFKTNPDF
NNHDELLADPRAYDDFVQECYAPAFLNKM AGGALLAGGGSTSCCYQQALEEIQAB
NSHLEQFCKDTGLGYRANGIVTAFDIMQRLENWYIDNGTLSFEESSNGKVKAIWQEQS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /product-"plasmid partition protein"
/protein_id-"BAB77490.1"
/db_xref-"GI:17134935"
/translation-"MTDSNKQVFVRFRVDEEKRDRFKITCIQLKTTMDEVLKGLLDKW
LEENDPESTKNK"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VDRADQDWHCNLESKLDINPNYDFENEIEGDKLVWEVCGEVPASTTADGTGVDIAVPG
BLIAMPESLLTTEVESQLKSVDVAEVCLQGEEEASHOTVNVQRDBQPKSSAFGSTPCS
TSGSTLELSFEQKLMGNVALIRECIADESWGLIAELIEEWTSEFKSAVWKELTPEERQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /transl_table=11
/protein_id="BAB77491.1"
/db_xref="G134936"
/translation="MTDLAEMPENALPRHQGTLTLIAGYDLGNSGVKFVTSDRKIRFP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KVTQSLKHLLAALSNYPYKPVINLIICASLHERGDLEEQLIDAIAGKHIVKFGGKPIP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TTVNIHVLKVYDEGHAAIAANAHTLDTSKQNVIVDIGNRTVIATLIGQKGHLANRKTF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DNGVEELIRMISVNPTFKNRLYGEIAIPHLIRQGLESSEKPFWYGRQFSFEDVYRQEL
MPWYQKSLAPVFKFIHPWKINADACLIIGGGSQLPSVDEALKAKGFVIAENPLWANAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KPSDKTVKAINQALPRIKALFPKTKLVTVPHPSGKRKLQALQGISFDNVPMSITGTPI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A I SENSTPNSTPNPHQQSLINQDFHTTHTSDEGLDEKNNSGFQKVDLDQQPKSEINVN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SYLENCYYRPTELPTEGYVEYLEGDAITKLDYKQWLSGYAAYDANPKNHLRVTDDATA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /transl_table=11
/protein_id="BAB77492.1"
/db_xref="G1:17134937"
                                                                                                           /gene="al19003"
complement(4726. .8019)
/gene="al19003"
/note="ORF_ID:al19003
                                                                                    complement(4726. .8019)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            complement(8400. .8570)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   complement(8400. .8570)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /gene="parG"
/note="ORF_ID:as19004
parG gene product"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note="ORF_ID:alr9005
unknown protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /note-"ORF_ID:asr9006
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8828. .9880
/gene="alr9005"
8828. .9880
/gene="alr9005"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /codon_start=1
/transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GLYQLATMMYSRGIDE
                                                                                                                                                                                                                         unknown protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     unknown protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9921. .10088
/gene="asr9006"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9921. .10088
/gene="asr9006"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /gene="parg"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AVKQLKALE"
                                                                                    gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   gene
                                                                                                                                    CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /translation="invalaticnvekiksfonvggsekhtarlodtpnadttkknirli
Gmednssledluktkianttkhkprkdavlcsemflsaspeyfrdddsragewher
Wylfrgaskowlinnygdkcyraelhideatphihativpindrtkolshkemegog
Wylfrgaskowlinnygdkcyraelhideatphihativpindktkolshkemegog
Wylfrgaskopstasaldierbykgskathtykeyyqavnsepliavwsnkl
Apqplesathvyariondoghalinholadbramgerleraedraksekerorleke
Vrslelktqolrdlelldvamelgldysrermrghghiinindgskfydfspdockgg
                                                                                                                                                                                                                   circular BCT 05-DEC-2001 DNA, complete genome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GAIDLVMHVNQCNFRQAVVWLHERFGEAGVERAAIAHYKNRAADIIQAEPRPOFTPPV
ENKANWPAVERYLTQQRGIPSDCLQMLQNLGLLYADDQQNAVFVMRDLEGQTSGAFLR
GTRGENNSFKGYSKGTKRSDSWFYFGLGGQATDKVAHVLLCSSPIEAVSRAMLEYFVR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GNVPPERTLYMAVDNINSLPVERLQNVPNÍLVTFGKDQSTNAAAQRVLELLPQSQQVL
SKAADWNQQLLEYGWQLRRQQYQQODDELSI"
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Kaneko,T., Nakamura,Y., Wolk,C.P., Kuritz,T., Sasamoto,S., Watanabe,A., Iriguchi,M., Ishikawa,A., Kawashima,K., Kimura,T., Kishida,Y., Kohara,M., Matsumoto,M., Matsumo,A., Muraki,A., Nakazaki,N., Shimpo,S., Sugimoto,M., Takazawa,M., Yamada,M., Yasuda,M. and Tabata,S. Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium Anabaena sp. strain PCC 7120
DNA Res. 8 (5), 205-213 (2001)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Submitted (02-MAY-2001) Takakazu Kaneko, Kazusa DNA Research Institute, The First Laboratory for Plant Gene Research; Yana 1532-3, Kisarazu, Chiba 292-0812, Japan (E-mail:kaneko@kazusa.or.jp, URL:http://www.kazusa.or.jp/cyanobase/, Tel:81-438-52-3935(ex.2338), Fax:81-438-52-3934)
                                                                                                                                                                                                                                                                                                                                                                                  Nostoc sp. PCC 7120
Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
                                                                                                                                                                                                                                                                                                                                                     Nostoc sp. PCC 7120 plasmid:pCC7120epsilon DNA.
Nostoc sp. PCC 7120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1. .40340
/organism="Nostoc sp. PCC 7120"
/plasmid="PCC7120epsilon"
/db_xref="taxon:103690"
/note="synonym:Anabaena sp. PCC7120"
                                                                                                                                                                                                  Nostoc sp. PCC 7120 plasmid pCC7120epsilon AP003605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /gene="alr9001"
/note="ORF_ID:alr9001
similar to mobilization protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /transl_table=11
/protein_id="BAB77487.1"
/db_xref="GI:17134932"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /transl_table=11
/protein_id="BAB77488.1"
/db_xref="G1:17134933"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          complement(3399. .3833)
/gene="al19002"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                complement(3399. .3833)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /gene="all9002"
/note="ORF_ID:all9002
unknown protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1429. .3222
/gene="alr9001"
                                                                                                                                                                                                                                                                                                AP003605.1 GI:17134931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2 (bases 1 to 40340)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Direct Submission
                                                        1429
                      Oy 2647 gcgccgggt 2655
                                                                                                                                                                                                                                                                                                                                                                                                                                          (sites)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              11759840
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Kanèko, T
                                                                                                                                                                                                                                       DEFINITION
ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                dene
                                                                                                                                                           RESULT 1:
AP003605
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE
                                                                                                                                                                                                                                                                                                                                                        SOURCE
```

```
Rattus,

Rat
                                                                                                                                                                                                                                                                                                                                                                                                                                       ACUYD146
183544 bp DNA linear HTG 20-DEC-2001
Rattus norvegicus clone CH230-8G23, *** SEQUENCING IN PROGRESS ***, 68 unordered pieces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Métazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                                                                                                                                                                                              15366 CGCTGGAGCCGGAGGCGTCGCTTTCGCCGTCGCCGGAGCCGGTTGCCGTTGCCGTTGCCGT 15425
                                                                                                                                                                 2528 egecgagecegtettegteacegagecegtettegteacegtettegtegeegageecaa 2587
                                                   DD 15546 AGCCTTCGCCGTCGCTGTCGCCG 15568
                                                                                                                                                                                                                                                                         2588 egecgtegtegtegeeggtgteg 2610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AC095146.2 GI:17942112
HTG; HTGS_PHASE1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (bases 1 to 183544)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Rattus norvegicus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Worley, K.C.
Direct Submission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Norway rat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AC095146
                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 13
AC095146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AUTHORS
  a
                                                                                                             qq
                                                                                                                                                                                                                     q
                                                                                                                                                              ò
                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /transl_table=11
/protein_id="pab77494.1"
/protein_id="pab77494.1"
/protein_id="pab77494.1"
/protein_id="pab77494.1"
/translation="wawOpPLOELVAQAQELLSAIRQHPQYQALQIDADVTLGDVTQFF
NTLQWSATSDAVDTTREGFFQ"
NTLQWSATSDAVDTTREGFFQ"
/gene="asr9009"
10654..10824
                                                                                                                                                                                                                                                                                                     /translation="wanningvidstilstrkvvQidrtlvQvlvQTGTiQAPQFIFRP
LAGQRKKADLKLNHKALTNRCYEVEGATTKSSVLSQESLQLSIF"
10372. .10569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /protein_id="Bab77496.1"
/db_xref="G1:17134941"
/db_xref="G1:17134941"
/translation="MDHGIAVINPVAGHSSHIDEKAQTWGGFFADILGYLGDINKLEQ
/translation="MDHGIAVINPVAGHSSHIDEKAQTWGGFFADILGYLGDINKLEQ
PAPARNNALELAERLDPFLENARSAFEALQKLTTGQYTWTELRKQYGSHVANAINKIR
VLNAQFDSEWQQLDAFDRAELLKIDQKRQHGLAEIAAQLHVDLQAELERHQNKLTGIF
NRQQYQTERQTITAGLREKRQALLNRAKYGSTGNDQQPREQIPVTVQASPAASSVSAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /transl_table=11
fprotein_td="pab77495.1"
/db_xref="GI:134940"
/translation="MDSKYYLCEAENTDEGVNKVTPYDKPEEALEAASNSTAKVHFIS
/translation="MLKQLSEQLNIDDYGELVNYILLEFQRRGGEFSESVEMPSKKVE
SSLANKLAAMF"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Db 15306 caccarcaccarcacrarradacasaccasasaccarcacrircaccaraccrircacrir 15365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2408 gcccgagcccgagcagctcgccatcgccgtcgccgtcgccgagctcgaggtccatcctcgt 2467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2348 cyccytccgyatcyccytctccytcygtgaytccytcygcygattcyccgaycctgaycccya 2407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .
,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
3.1%; Score 90.2; DB 1; Length 40340;
Best Local Similarity 58.9%; Pred. No. 0.077;
Matches 155; Conservative 0; Mismatches 108; Indels 0;
                                                                                                                                                                                                                                                                         /protein_id="BAB77493.1"
/db_xref="GI:17134938"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /transl_table=11
/protein_id="BAB77497.1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10824. .11540
/gene="alr9010"
10824. .11540
/gene="alr9010"
/note="ORF_ID:alr9010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /note="ORF_ID:asr9008
unknown protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /note="ORF_ID:asr9009
unknown protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note="ORF_ID:alr9011
                                                                                                                                                        /note="ORF_ID:asr9007
unknown protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GTVRGFGSFLGNLWQRLGDR
                                                                                                                                                                                                                                                                                                                                                                                                          /gene="asr9008"
10372. .10569
/gene="asr9008"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             unknown protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       unknown protein"
                                                10094. .10360
/gene="asr9007"
10094. .10360
                                                                                                                                                                                                                                                  /transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /gene="asr9009"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /codon_start=1
/transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11574. .12113
/gene="alr9011"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11574. .12113
/gene="alr9011"
                                                                                                                                 /gene="asr9007"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     codon_start=1
                                                                                                                                                                                                                        /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "VNPLAVDEEEE"
                                                   gene '
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      gene
                                                                                                             CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
Length 183544;
          bp in length
length
                              bp in length
length
                                                                                                                                                                                                       bp in length
length
                                                                                                                                                                                                                           bp in length
length
                                                                                                                                                                                                                                                                                                                                                                                                   bp in length
length
                                                  bp in length
                                                                      bp in length
                                                                                         length
                                                                                                           bp in length
                                                                                                                             bp in length
                                                                                                                                                 bp in length
                                                                                                                                                                    length
                                                                                                                                                                                      length
                                                                                                                                                                                                                                              bp in length
                                                                                                                                                                                                                                                                 bp in length
                                                                                                                                                                                                                                                                                     bp in length
                                                                                                                                                                                                                                                                                                        in length
                                                                                                                                                                                                                                                                                                                          bp in length
                                                                                                                                                                                                                                                                                                                                           in length
                                                                                                                                                                                                                                                                                                                                                                length
                                                                                                                                                                                                                                                                                                                                                                                  length
                                                                                                                                                                                                                                                                                                                                                                                                                       bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                         bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                            bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 bp in length
length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            bp in length
length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            in.length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          bp in length
                                                            length
                                                                              Length
                                                                                                 length
                                                                                                                    length
                                                                                                                                        length
                                                                                                                                                         length
                                                                                                                                                                           length
                                                                                                                                                                                              length
                                                                                                                                                                                                                                                        length
                                                                                                                                                                                                                                                                         length
                                                                                                                                                                                                                                                                                             ength.
                                                                                                                                                                                                                                                                                                                length
                                                                                                                                                                                                                                                                                                                                   length
                                                                                                                                                                                                                                                                                                                                                       ength
                                                                                                                                                                                                                                                                                                                                                                        Length
                                                                                                                                                                                                                                                                                                                                                                                           length
                                                                                                                                                                                                                                                                                                                                                                                                                                length
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          length
                                                                                                                                                                                       pp in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 2;
                                                                              unknown
of 2568
                                                                                                                                                                                                                                                        unknown
of 1495
                                                                                                         2746
                                                                                                                            3410
                                                                                                                                                4346
                                                                                                                                                                                                                                                                                                                         1156
                                                                                                                                                                                                                                                                                                                                           1345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1595
                                                                                                                                                                                                                                                                         unknown
                                                                                                                                                                                                                                                                                                                                                                       unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                            2065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1318
                                                                                                                    unknown
                                                                                                                                                                   1948
                                                                                                                                                                          unknown
                                                                                                                                                                                                                           of 1352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1916
                                                                                                                                        unknown
                                                                                                                                                                                                                                                                                                                 unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1958
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 86.4;
                                                                                                                                                                                                                          contig
gap of
contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                      contig
gap of
contig
gap of
                              contig
gap of
contig
                                                                                                gap of
contig
gap of
                                                                                                                                                                                                                                                                                   contig
gap of
contig
gap of
                                                                                                                                                                                                                                                                                                                         contig
gap of
                                                                                                                                                                                                                                                                                                                                                              contig
gap of
contig
gap of
contig
            contig
gap of
                                                                              gap of
contig
                                                                                                                             contig
gap of
                                                                                                                                               contig
gap of
                                                                                                                                                                   contig
gap of
                                                                                                                                                                                              gap of
contig
gap of
                                                                                                                                                                                                                                                                contig
gap of
                                                                                                                                                                                                                                                                                                                                           contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         gap of
contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      gap of
                                                          gap of
                                                                                                                                                                                      contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                            contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       contig
                                                                      contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        contiq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gap of
           94688:
94788:
97736:
97836:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      166131:
 92457
92557
94689
94789
97737
97837
100284
                                                                              103336
103436
103436
1061004
108850
108950
112360
112806
116906
118854
1118954
1118954
1118954
1118151181
                                                                                                                                                                                                                          123773
125125
125225
126339
126439
127934
128034
129917
                                                                                                                                                                                                                                                                                                                132152
132252
133408
133508
134853
134953
137210
137310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    145553
145653
147297
147397
148992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    149092
150399
150499
152457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           152557
154952
155052
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           164714
                                                                                                                                                                                                                                                                                                                                                                                                  138527
139724
139824
                                                                                                                                                                                                                                                                                                                                                                                                                               141818
                                                                                                                                                                                                                                                                                                                                                                                                                                         41918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          59195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       56914
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                              NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
NOTE: This is a "working draft' sequence. It currently consists of 68 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as arbitrary. Gaps between the contigs are represented as This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
                                                                                                                                             Consensus quality: 148905 bases at least Q40 Consensus quality: 156308 bases at least Q30 Consensus quality: 166905 bases at least Q20 Estimated insert size: 150335; sum-of-contigs estimation Quality coverage: 0x in Q20 bases; agarose-fp estimation Quality coverage: 2.2x in Q20 bases; sum-of-contigs estimation
Submitted (16-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
Dec 20, 2001 His sequence version replaced gi:15625700.
Center: Baylor College of Medicine
                                                                                                                             to
                                                        10936: contig of 10936 bp in length
11036: gap of unknown length
17011: contig of 5975 bp in length
                                                                                                                                                                                                                                                                                                                                                              unknown length '
of 5348 bp in length
unknown length
of 5631 bp in length
                                                                                                                                                                                                                                                                                                                                          gap of unknown length
contig of 4708 bp in length
                                                                                                                                                                                                                                                                                                                                                                               length
bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                               bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                  bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 length
bp in length
                                                                                                                                                                                                                                                                                                                                                                                                             in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          in length
                                                                                                                                                                                                                                                                                                                                                                                                  length
                                                                                                                                                                                                                                                                                                                                                                                                                       length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              length
                                                                                                                                                                                                                                                                                                                                                                                                                                         length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            length
                                                                                                                                                                                                                                                                                                                                                                                                                                                            length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              unknown leng
of 2642 bp
                                                                                                                                                                                                                                                                                                                                                                                                   unknown l
of 5140 b
                                                                                                                                                                                                                                                                                                                                                                                                                                         unknown J
of 5666 h
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              unknown
of 4108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 unknown
of 3263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     unknown
of 3797
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             unknown
of 3867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2495
                                                                                                                                                                                                                                                                                                                                                                                                                               of 4523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                of 3711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   of 2581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οĘ
                                                                                                                                                                                                                                                                                                                                                                       contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             gap of
contiq
                                                                                                                                                                                                                                                                                                                                                                                           contig
                                                                                                                                                                                                                                                                                                                                                                                                                                contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                   contiq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         contiq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                contig
                                                                                                                                                                                                                                                                                                                                                                                                                       of
                                                                                                                                                                                                                                                                                                                                                                                                                                                           of,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οĮ
                                                                                                                                                                                                                                                                                                                                                                gap of
                                                                                                                                                                                                                                                                                                                                                                                 gap of
                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                          ō
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ō
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ō,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           o
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ō
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      gap of
                                                                                                                                                                                                                                                                                                                                           7111:
1819:
1919:
7267:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  70511:
70611:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              67830:
67930:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               74578:
                                                                                                                                                                                                                                                                                                                                                                                                                               12861:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 56859:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             54019:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            30928:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92456:
                                                                                                                                                                                                                                                                                                                                                     171112
221920
277268
332099
3330099
3330099
3330099
3330099
332099
332099
332099
332099
34209
3520523
442050
442050
442050
442050
442050
442050
442050
442050
442050
442050
442050
442050
442050
442050
442050
442050
442050
                                                                                                                                                                                                                                                                                                                        10937
                                                                                                                                                                                                                                                                                                                                           7012
  JOURNAL
                              COMMENT
```

g

ò

염

δ

```
18545:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   21984:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   34213:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13240:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        32510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  38418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      38518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        30361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        32610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    30461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          be preserved.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   findPhrapList
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24424
255965
226066
228008
330362
330462
330462
330462
33131
34214
34214
34214
34214
34214
34214
34214
34214
34214
34214
34214
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
34219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5463
5563
11132
111232
114472
14572
18546
19690
19890
21985
22085
   TITLE
JOURNAL
                                                                                                                                                          COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Muzny, D.M., Adams, C., Adio-Oduola, B., All-osman, F.R., Allen, C., Ashans, C., Adio-Oduola, B., All-osman, F.R., Allen, C., Ababrooks, S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barbaria, J., Barton, J.E., Burch, P., Burkett, C., Burcall, K.L., Brydan, N.P., Carron, T.E., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Carron, T.E., Carter, M., Dadhory, I., Charistopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D. Dathorne, S.R., Delaney, R.R., Delandy, C., Davy-Carroll, L., Dederlich, D.A., Delaney, R.R., Delandy, C., Davy-Carroll, L., Dederlich, D.A., Delaney, R.R., Delandy, C., Davy-Carroll, L., Dederlich, D.A., Delaney, R.R., Delandy, C., Edgar, D., Edward, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisl, A., Gao, J., Garner, T., Garza, N., Gall, R., Gorrer, M., Garrer, M., Garrer, M., Gall, R., Gorrer, M., Garrer, M., Jackson, E., Hernandez, J., Hernandez, O., Hodgson, A., Hollows, M., Hollows, J., Jackson, E., Karlsson, E., Howard, S., Khalsson, E., Man, M., King, L., Karlsson, E., Man, M., Karls, M., Karlsson, E., Man, M., Karls, M., Marlsson, E., Man, Marls, A., Marlsson, E., Man, Massel, M., Ren, Y., Rives, M., Rojus, M., Ren, Y., Rives, M., Rojus, M., Pan, S., Pace, A., Tamerisa, R., Tang, H., Tang, M., Tang, M
                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PACUY941/
8Attus norvegicus clone CH230-2M18, *** SEQUENCING IN PROGRESS ***, 51 unordered pieces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Chordata; Craniata; Vertebrata; Euteleostomi;
                                                                                                                                                                                                                                                                                                         73792 GCTTCTCTCCGTCTCCGTCTCCGTCTCCGTCTCCGTCTCCGTCTCCGTCTCCGTCTCCGTCTCCGT 73851
                                                                                                                                                                                                                                                                     2480 cgccgtcaccatcgccgagtccgtctcggtctccgtcaccatcggcgtcgccgagcccgt 2539
                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Rodentia; Sciurognathi; Muridae; Murinae;
                                                                                                           2420 gcagotogccatogccgtcgccgagotogagotogagtccatoctcgtcgccgtctccgt
                                           ;
0
                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                73912 CTCCGTCTCTCCCTCTCCCTCTCCCTCTCCC 73951
                                                                                                                                                                                                                                                                                                                                                                                                                    2540 cttcgtcaccgagcccgtcttcgtcaccgtcttcgtcgcc 2579
                                           46;
   Pred. No. 0.14;
                                       0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AC099417.2 GI:17940849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (bases 1 to 97225)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unpublished
2 (bases 1 to 97225)
71.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Metazoa;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mammalia; Eutheria;
                                           114; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HTG; HTGS_PHASE1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Norway rat.
Rattus norvegicus
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Rattus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JOURNAL
REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AC099417
```

2 (Dases . Worley, K.C.

AUTHORS

```
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently consists of 51 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as urns of N. but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will
Direct Submission
Submitted (15-NOV-2001) Human Genome Sequencing Center, Department
Of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:16930824.
                                                                                                                                                                                                                                    Consensus quality: 74561 bases at least Q40
Consensus quality: 81622 bases at least Q30
Consensus quality: 88269 bases at least Q30
Estimated insert size: 71621; sum-of-contigs estimation
Quality coverage: Ox in Q20 bases; agarose-fp estimation
Quality coverage: O.8x in Q20 bases; sum-of-contigs estimation
                                                                                                                                                                                                                                                                                                                                               NOTE: Estimated insert size may differ from sequence length
                                                                                                                                                                                             ------ Summary Statistics
Assembly program: Phrap; version 0.990329First call to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              length
bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          unknown length
of 3240 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  contig of 1144 bp in length
gap of unknown length
contig of 2095 bp in length
gap of unknown length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        bp in length
length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ength.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         length
                                                                                                                  length
                                                                                     Center: Baylor College of Medicine
Center code: BCM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             contig of 5462 b
gap of unknown 1
contig of 5569 b
gap of unknown 1
contig of 3240 b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2816
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                of 1910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1842
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       of 3974
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                of 2239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             of 1541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         of 1943
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2049
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         of 1189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       of 1438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       of 1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              of 1839
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  oţ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ot
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             oţ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      gap of contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            contig
gap of
contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gap of
contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       gap of
contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               41331
43241
43341
45180
47145
47245
49087
```

```
/codon_start=1
/product="Pol"
/product="Pol"
/protein_id="Aak85565.1"
/protein_id="Aak85565.1"
/db_xref="G1:1213126"
/translation="MPDYSYRPTIGRTYVYDNKYYKNLGSVIKNAKRKKHLLEHEEDE
/translation="MPDYSYRPTIGRTYVYDNKYYKNDTWKLYVWSGREFLRETWTR
VEDSFPIVNDGEWMDYFLVINLRPTRPNRCYKFLAQHALRWDCDYVPHEVIRIVEPSY
VGMNNEYRISLAKKGGGCPINNIHSBYTNSFESFVNRVIWENFYKPIVYIGFDSGEEE
EILIEVSLVFKVKEFRPDAADFTGFAY"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /protein_id="Aak85567.1"
/db_xref="G1:15213128"
/tb_xref="G1:15213128"
/translation="MGASQDSEQKIIKLFYCWSSQTGAGLDNDKDLHCLYDLERFVGAHRRQADEKAKKKKSAEKALKRYEIAADRHMLEAAATPACTDDGRWSTLSQAQLDDIAREKIYHLQLKQDRLLKQ"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AY043265 118584 bp DNA circular VRL 19-AUG-2001
Epiphysa postvittana nucleopolyhedrovirus, complete genome.
AY043265.1 GI:15213125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /translation="MQFELREIYDNYLQRATAPIMCGTCSETMNVSQYPEHLTTVHKK CPLQLSCVRCDDQSSWGQFSASGLMFEHMYFCLQNYVQNGV" complement(1713. .2090)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Box
                                                                     1 (bases 1 to 118584)
Hyink,O., Dellow,R.A., Olsen,M., Caradoc-Davies,K.M.B., Drake,K.
and Ward,V.K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hydroc., Delow, R.A., Olsen, M., Caradoc-Davies, K.M.B., Drake, K. and Ward, V.K. and Ward, V.K. Driect Submitsion

Submitted (03-JUL-2001) Microbiology, University of Otago, PO Bo. 56, Dunedin, New Zealand

Location/Qualifiers
                                  2424 ctcgccatcgccgtcgccgtcgccgagctcgagtccatcctcgtcgccgtctccgtcgcc 2483
                                                                                                                                                        gicaccatogoogagicogictoogictoogicaccatoggogicgoogagooogictic 2543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1. .118584
/organism="Epiphyas postvittana nucleopolyhedrovirus"
/db_xref="taxon:70600"
                                                                                                                                                                                                                                                                          Epiphyas postvittana nucleopolyhedrovirus
Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
Nucleopolyhedrovirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The complete sequence of the Epiphyas postvittana nucleopolyhedrovirus genome Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Epiphyas postvittana nucleopolyhedrovirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /product="unknown"
/protein_id="AAK85566.1"
/db_xref="G1:15213127"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       complement(606. .1343)
/gene="pol"
/note="ORF1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       complement(1434. .1691)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 complement(606. .1343)
/gene="pol"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /product="unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /note="EppoMNPV"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (bases 1 to 118584)
                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 15
AY043265/c
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM
                                                                                                                                                        2484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOURCE
                                                                                           a
                                                                                                                                                        δ
                                                                                                                                                                                                              Dp
                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5059 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      unknown length
of 1345 bp in length
unknown length
                                                                                                                                                                                                                                                bp in length
                                                                                                                                                                                                                                                                                                                                                                   bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           length
bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 of 1136 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               length
                                                                                                                                                                                                                                                                                                                                                                                                                               in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        of 1458 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1109.bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        unknown length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                unknown length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           contig of 1716 bp in 1
gap of unknown length
contig of 1379 bp in 1
gap of unknown length
                                                                                               length
                                                                                                                                                           length
                                                                                                                                                                                                                    length
                                                                                                                                                                                                                                                                                length
                                                                                                                                                                                                                                                                                                                                      length
                                                                                                                                                                                                                                                                                                                                                                                                                                                               length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     unknown length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 unknown length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ength
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                contig of 1350 bp
gap of unknown len
contig of 1513 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  of 1469 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              contig of 1264 bp

    .97225
/organism="Rattus norvegicus"

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 contig of 1136 b
gap of unknown 1
contig of 1381 b
gap of unknown 1
contig of 1294 b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 24036 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                           unknown
of 2288 N
unknown
of 1433 N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  unknown
of 1428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    unknown
of 1589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                unknown
of 1282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           unknown
of 1044
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      unknown
of 1617
                                                                                                                                                                                                                                                                          unknown
of 1734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              unknown
of 1022
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      of 1584
unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   of 1785
                                                                                                                                                                                                                                                                                                                                         unknown
                                                                                                                                                                                                                                                                                                                                                                                                                               of 1063
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    of 1357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /db_xref="taxon:10116"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /clone="CH230-2M18"
22043 c 22307 α
                                                                                                                     contig
gap of
contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   contig
gap of
gap of
contig
contig
gap of
contig
c
contig
c
contig
contig
contig
contig
contig
contig
contig
contig
conti
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  location/Qualifiers
                                                                                                                                                                                                                                                                                                          contig
gap of
                                                                                                                                                                                                                                                                                                                                                               contig
gap of
contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 contig
gap of
contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      contig
gap of
                                                                                                                                                                                                                                                contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gap of
                                                                                                                                                                                                              gap of
                                                                                                                                                                                                                                                                                                                                                               60919:
61019:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 64570:
66003:
                                                                                                                                                                                                                                                                                                                                         58861:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    73880:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    75569:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           81062:
$50298
$50298
$50398
$52009
$52109
$53251
$53351
$5536
$5536
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$55036
$550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 664471.
664571.
664671.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.
6661.

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23780
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURES
```

ó

Gaps

0;

Query Match 7.2%; Score 84.2; DB 2; Length 97225; Best Local Similarity 67.2%; Pred. No. 0.29; Matches 119; Conservative 0; Mismatches 58; Indels 0;

```
LLKIOFGPTTPFTIRELRNNVOLLLINIHPVYDNNRPVPPSVOYLGGGHILTOTOPDOL
OKTLERQLNASVNGTYVSFGSSIOPKSIHDEFLQMLVNTFTGLNNRTVLWKVDDAVV
DSKLLPPWTTONWFNQRAVLHHKNTVFYTGGGGLOSDEBALHAQVPWYCLPWMGDQF
HHANKLQERVARTLHTVDVSASOLILAITDVIVNKITYQTRWAELRAVIDYDEIAPA
DKAIKFTERVIKKFGHDITYPARSLKSPAANLDQSNYFISFPL"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /codon_start=1
/poduct="unknown"
/product="unknown"
/protein_id="ARR85577.1"
/db_xref="G1:15213138"
/translation="METPKMLPSSFATPKRAALGTFVKTVYTTTVSGRGLTRDERNR
IFVAYAQLKKTRLSFRILTORSKRYKRNNITADLAAQLERERGESKGSK
YFAVICONGVLITISGSGGFVRORVANMCAVGGGIFCERRNDCARDROLIAEALAAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FAKFKULISITNENNEHLQATVCSQYDSVCIMNPCSQQEIIFDGFIKHDDEGVTVPFI
VGPLFSLCESVWAPHYRATVDAIERLQTVLKYFINBAYLKGRATTLKKLLFSDNNESD
LVNNIVNFINVDKIECSEKYVNVTKWVPAINYVTGKQLLTVLFIFKFN"
complement(10162. .11235)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AFDNITQLTANITROPELTILKCCPNGAAYCFGNDDNILELLEHIDTRATHFUNEFDLI
RFALQVEMWTQNGYDEYTILKCCPNGAAYCFGNDDNILELLEHIDTRATHFUNEFDLI
RFALQVEMWTQNGYDEYNNQKLMINKRHAVIENVFFVNYRWHPFFGERCSKKFINF
GDNIHVLVSPHRVLMDQIMCLLKDIFTDKLEFKVBRRKTLIRALFAKLPQDACDVCV
NNHSDVAVGKHRFENITSFCKKTLDIYGPALGCRKLVYAYLTTNTFLNQIPHYVADYA
NYPHNROSCEQXWREFMGYMYLL"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /trasiatation="MHFAVILTLLAAQAHAANILAVLPTPAYSHHLVYKAYVQALANK
CHNYYVIKOLLNYVDSHAHGGHIGO,TADMASBQYKKLVSSSGAFKKGVVADETT
VTANYVIKGLIDMFKDOFNNYVQCFLLCNTFDVVVVEAFADVALGHIERPAPVIO
IAPGYGLDENFDTAGAVARHPKYYPNIWRSTFVGGTAGALSEWRLYNEFELLARYSDK
                                                                                                                                                                                                                                                                                                                                                                       /translation="MATYMYTLDRVNLAWNSIAYNNSRKFAFWTLKQRWIHVDRHFDN
AQLLNETNSNSVSDYHVKPLDEGGREWVIDAFRKNYVNKFELMLKINVGATALLLEF
TEMHYHRINFSGNRGFPMMYLLAGGFRIGALKSVREHWYRIFERFRLAMRDIRNGSF
AHCMQRAVNNYQDQLILGSSKPDELILHFWPDVDKDVFCNPTKQIRAPFSYNHKGTQFS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /product="unknown"
/protein_id="AAK85578.1"
/d_xref="di:15213139"
/translation="MDCTITLIPISLKNIEEPKRSECFKLTSMREDAEFCLNVKCRSP
                          KTVLKHIKTLNVNSDKFINAHKLFENQVCARFEQLEQRLETLERVPDAPTMPGVIFPR
DVNKHQHLAVFVNQERGNTQIGFARGQEEYFRKRKLEFEEEDMHKMLETVHPNPQMAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /translation="MDCVANQVCSGTLPYVNTRDLEDCLRNKITAKIGLRFFRDCFEA
VTADKSGLFVLSGGAAAACHIDEKRDMLKCLDFDYYNCSKEWLQLAQMQKHLQACVDV
VFKTPPVEISPDSMGTVYATKHGLIQILQQLSFEYKDDVLLAIKTDKGYDCDDVRDNI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 83.6; DB 14; Length 118584; Pred. No. 0.32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LGSDVITSAANKRFKILNVEKVVNAKFIVQQTLHNGFNNYPNTH
                                                                                   QCIKDRFISNGYKIKKVSNRRRVIEVDCNINAAKDIVNNVIIS"
complement(6427. .7182)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RCITQOLLEKIONCSTGCSDGGASTSSISNITH"
7254. 8732
/gene="egt"
/gene="egt"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /protein_id="AAK85576.1"
/db_xref="G1:15213137"
                                                                                                                                                                                                                                                                                                              /protein_id="AAK85575.1"
/db_xref="G1:15213136"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /protein_id="AAK85579.1"
                                                                                                                                                                    complement(6427. .7182)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /db_xref="GI:15213140"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /product="unknown"
                                                                                                                                                                                                                                                                                      /product="LEF-1"
                                                                                                                                                                                                                                                                 /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /product="EGT"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8914. .9528
/note="ORF13"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9497. .10123
/note="ORF14"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note="OKF12"
                                                                                                                                         /gene="lef-1"
                                                                                                                                                                                                     /gene="lef-1"
                                                                                                                                                                                                                             /note="ORF11"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match .
Best Local Similarity
                                                                                                                 gene
                                                                                                                                                                          CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /protein_id="AAK85570.1"
/db_xref="G1:15213131"
/translation="MADARFLRFSTRLTEEYKENVVAQIDHLTNLRALINAKTTIANV
RRFGFVDRNAFVAACWAVIVHTYASSDDKTTLQPEGLYFRVCRFCNEIADVFDDDHS
ITRYLCAACGTCLVVDNPLNVFTEDGVQRFMEIQRINAGGEP"
                                                                                                                                                                                                                                                                                      QLVEISGMRLFALLTMSKPTEPKRQFARRSKRNICLKACADGSQTLHKSLTGLKLPLC
IAKIMAELSNASVPRGGMYRKRFEFNCYLGNVLTCTKCKSACLIGALLHFYKMDAKCV
GEVTHLLIKSQDVYKPSNCAKMKSMKLCPQANMCKGLNPICNF"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /translation="MFPDRWHDYTVCGNVIEGTNFICFKVPLREEMFEYVTNDLDRWT
IKNLIKQQCSLGAVIDLTNTFRYYDNANMRDEGLLYKKIRVPGQVLPDENVVQLFFDV
VKHFTARCPGMLIGVHCTHGLNRTGYLVCRYMINILNISPQDAIARFETARGHKIERQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ĎVÞWMHREIDTARQILSGALCVNDACHMLDLDNTKERLATWYRCGNTKGVCHEVYVVL
KAIDALVPIDKRSGGQFVDVDTFDTILLNITNKLPEVLHQYLRFVYYQKLEHVTDVFN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PKSDTVGWWYNKFCVLTYMYRIINKSVPAELLTRLØKAVVETIOPNAGDKYNCAEVMA
VYGRECGISKQHFABHKTANNHIMFKSMRGØTTMVDBRYKSFSVIKDFSRHCKETYL
DLKSHVDSLYINGTTDKØKNAMFDLLCCQSEQDLDVDCYDYIVNMLFFK"
complement (5163 . 5516)
                                                                                                                                                                                                                                                              /translation="mDNVAVWNPSANTATIKKNETYLIDPNDFMDVLTLSPYTVFRRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /translation="MIYLRYYIFFVFFRESRTKSKKPRISHIKLFLFYATLHYNMNRR
INNAPVMVAGHDYDREQLKRDINSLRHSVHELCKRSTTGFDCNRLLESSTDTAKPTVV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /translation="MSLSSKLLVYDYYGNSMCGVNEQYGASYHLYKIVDEHITNSYVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /translation="MSSNNSFMDVDNTHENVETNVKDTLLKPINVWSNDPSSQARCEK
GAPPVYFAYKCVPLESPLRGNQYKIDPTLPALQKITTDTANKTPVLASVKPNNYQDKN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /translation="MFDWMFGWWSFNEQHQQHHPLANHFNAQDVKQYAVDRRAQSDLV
NRSVFKCHPFTFKFRYVIDDDNGRCCRVVDFCKGLEINHELMLNCKWDSKHVRHLNEI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        'note="ORF9; unique to EppoMNPV"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /protein_id="AAK85574.1"
/db_xref="GI:15213135"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /protein_id="AAK85569.1"
/db_xref="GI:15213130"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /protein_id="AAK85572.1"
/db_xref="GI:15213133"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /protein_id="AAK85573.1"
/db_xref="GI:15213134"
                                                                                                                                                                                               /protein_id="AAK85568.1"
/db_xref="G1:15213129"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /protein_id="AAK85571.1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         complement(5531. .6490)
complement(2104. .2715)
                                                                                                                                                                                                                                                                                                                                                                       complement(2717. .3070)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     complement(3426. .3935)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           complement(3426. .3935)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   db_xref="GI:15213132"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /codon_start=1
/product="unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /product="unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /product="unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /product="unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /product="unknown"
                                                     complement(2104.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TVAEAYNNVLKNNII"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /product="PTP-1"
                                                                                                                                                                          /product="LEF-2"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IKTTAAEQHTICDKL"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                      /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /codon_start=1
                                                                                                                                               /codon_start=1
                                                                                   'qene="lef-2"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /gene="ptp-1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       note="ORF10"
                          /gene="lef-2'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note="ORF6"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      'gene="ptp-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /note="ORF7"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4168. .5145
/note="ORF8"
                                                                                                           /note="ORF4"
                                                                                                                                                                                                                                                                                                                                                                                                           /note="ORF5"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2990. .3424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NY IQHLLTL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        gene
  gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CDS
                                                        CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CDS
```

ó

```
Mycobacterium tube
Mycobacterium tube
Randomising oligon
PCR primer for 5'
                                                                                Human METH1 relate
AC004449 CDNA clon
Mycobacterium spec
Mycobacterium spec
Mycobacterium spec
Mycobacterium spec
Mycobacterium the
Mycobacterium tube
Mycobacterium tube
Mycobacterium tube
Mycobacterium tube
Mycobacterium tube
                                                                                                                                                                                                                                                                                                                                             M. tuberculosis an M. tuberculosis re Corn p-hydroxyphen A. cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                           DNA sequence encod
Active cellulase p
Streptomyces albid
Nucleotide sequenc
Active cellulase p
Mycobacterium tube
Nucleotide sequenc
                                                                                                                                                                                                                                                           Sequence containin
                                                                                                                                                                                                                                                                                            Drosophila melanog
Drosophila melanog
                                      DNA encoding novel
                                                             Human immune syste
                                                                          Human qene requlat
                                                                                                                                                                                                Nocardiopsis sp. p
                                                                                                                                                                                                                                                                     Thymidylate syntha
                                                                                                                                                                                                                                                                                                                      M. tuberculosis im Mycobacterium tube
                                                                                                                                                                                                                                                                                   Human thymidylate
                                                Tumour suppressor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Systematic polypeptide evolution by reverse translation; SPERT; ligand binding; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   New method of systematic polypeptide evolution by reverse translation - by linking each polypeptide in sample mixt. to individualised mRNA allowing further synthesis of selected
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Randomising oligonucleotide used in SPERT mRNA prepn.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ALIGNMENTS
                                                                                                                                                                                                                                                                                 AAC91215
ABL16717
ABL16716
                                                                                                                      AAX34249
AAX34248
AAV64542
                                                                                                                                                                      AAZ19343
AAZ19131
                                                                                                                                                                                                         AA199683
                                                                                                                                                                                                                      AA199682
                                                                                                                                                                                                                                 AAQ21833
AAQ36859
                                                                                                                                                                                                                                                                                                                                AAV44430
AAZ19340
                                                                                                                                                                                                                                                                                                                                                                                             AAX55662
AAD26526
                                      AAS84956
                                                  AAS46648
                                                                                                             AAX34250
                                                                                                                                                           AAV44433
                                                                                                                                                                                                                                                           AAF76910
                                                                                                                                                                                                                                                                     AAF31109
                                                                                                                                                                                                                                                                                                                      AAV64539
                                                                                                                                                                                                                                                                                                                                                         AAZ19128
                                                                                                                                                                                                                                                                                                                                                                      AAV04490
                         AAA58471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BP.
 AAQ21833 standard; DNA; 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    90US-0561968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           91WO-US05463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               08-JUN-1992 (first entry)
11707
4403765
58857
                                                                                                                                                                                                          4403765
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (COLS ) UNIV OF COLORADO
                                                6510
6510
3 6510
3 8186
3 8186
1164
1165
494
494
                                                                                                                                                                                                                                                                                  18596
3543
13837
                                                                                                                                                                                                                                                                                                                                  407
407
407
1766
2293
                                      651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1992-080018/10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gold L, Tuerk C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-AUG-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   02-AUG-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           W09202536-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  20-FEB-1992
 Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAQ21833
                                                                                                                                                                                                                                                                                                                                                         66.
666.
666.
666.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAQ21833/c
                                                  0 0 0 0 0 0 0 0 0 0 0 0 0 0
                                                                                                                                                                                                                                                                                                                00000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2, 2002, 09:11:38 ; Search time 307.67 Seconds
(without alignments)
16010.100 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence containin
Arabidopsis thalia
Arabidopsis thalia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Propionibacterium
Arabidopsis thalia
Arabidopsis thalia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Randomising oligon
PCR primer for 5'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DNA sequence encod
                                                                                                                                                                                                                                                                                                                                                                                      SIDSI/gcgdata/geneseq/genesegn-embl/NA1981.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1981.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1981.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1981.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1985.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1986.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1986.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1989.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1989.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1989.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1991.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1992.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1992.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1992.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1993.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1999.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1999.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1999.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1999.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1999.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1999.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1999.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1999.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1998.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1998.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1998.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1998.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1998.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1998.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1998.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1998.DAT:*
SIDSI/gcgdata/geneseq/genesegn-embl/NA1998.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                          1 atggatcgttcggagaacat........ggcggacacctacctgcagn 2869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
                                                                                                                                                                                                                                                                                                                                                                    /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.
              4.5
Compugen Ltd
                                                                                                                                                                                                                                             hits satisfying chosen parameters:
                                                                                                                                                                                                                      1736436 seqs, 858457221 residues
            GenCore version
Copyright (c) 1993 - 2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUMMARIES
                                                                                                                                                                                                                                                                                                       Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                            OM nucleic - nucleic search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAC47999
AAC52204
AAS59522
AAC4444
AAC45005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAQ21833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAF76910
                                                                                                                                                                                 IDENTITY_NUC Gapop 10.0 , Gapext 1.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Π
                                                                                                                                                                                                                                                                                                                                                          N_Geneseq_032802:*
                                                                                                                                                                                                                                                                     seq length: 0 seq length: 2000000000
                                                                                                                                 US-09-917-376-2
2869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2011
37286
777
                                                                                   July
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               oŧ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100.8
100.8
100.8
89
89
85.4
85.4
85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score
                                                                                                                                               Perfect score:
                                                                                                                                                                                  Scoring table:
```

Total number Minimum DB Maximum DB

Database

450/000

Result Š. U O υU

Searched:

Sequence:

Run on:

us-09-917-376-2.rng

```
Synthetic.
        molecule
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 3
AAF76910/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
                                                                                                                                                       The sequence is that of an example randomising oligonucleotide which is used in the prepn. of mRNA encoding candidate polypeptides for the method of systematic polypeptide evolution by reverse translation (SPERT). The method provides a rapid way of isolating and identifying polypeptide ligands which bind to target mols. The polypeptide ligands can be used in e.g. assay methods, diagnostic procedures, cell sorting, as activators or inhibitors of target mol. function, as probes, as sequestering agents, drug delivery vehicles, modififers of hormone action and as catalysts. See also AAQ21830-Q21832.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           tegecgagecegicticgicacegagecegicticgicacegicticgicgeegageeea 2586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Systematic peptide evolution by reverse translation; SPERT; ligand; specific; inhibitors; probes; assay; cell sorting; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2347 gegeegicegategeegicteegieggieggieggeiteggeitegeegageeigageeeg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2407 agecegagecegageagetegeeategeegtegeegtegeegagetegagteeateeteg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   tegecegicteegicaccaicategecegagicegicteeggicteecaicegeeg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Systematic polypeptide evolution by reverse translation - used for prodn. of polypeptide ligand specific for desired target
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 390;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PCR primer for 5' fixed sequence contg. T7 promoter and RBS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   102; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 390 BP; 125 A; 126 C; 133 G; 6 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 100.8; DB 1
Pred. No. 5.3e-10;
0; Mismatches 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Tuerk C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Example; Page 55; 102pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Smith JD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 3.5%;
Best Local Similarity 61.4%;
Matches 162; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AA036859 standard; DNA; 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92WO-US00801
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            91US-0739055
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        UYRE-) UNIV RES CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pribnow D,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1993-076529/09
    polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  22-JUN-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                31-JAN-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-AUG-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO9303172-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                18-FEB-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAQ36859;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sold L,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAQ36859/c
\mathbb{F} \times 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
```

```
SPERT is used to select novel polypeptides that bind the antibody of the epitope commonly recognised by the antisera from autoimmune mice which are the fl progeny of a cross of NZB and NZW parents (Portanova et al., J. Immunol. 144, 4633, 1990). The known epitope consists of ca. 10 amino acids at the N-terminus of the histone HZB consists of ca. 10 amino acids at the N-terminus of the histone HZB protein. To make mRNA encoding candidate polypeptides a 5' fixed sequence composed of a T7 promoter sequence and a ribosome binding site which is recognised by both prokaryotic and eukaryotic ribosomes, terminating in a restriction endouvelease site is synthesised and cloned using a number of oligonucleotides (example shown). A 3' fixed sequence is placed into a restriction site to provide an mRNA encoding the C-terminal trailer sequence of ca. 100 nucleotides lacking stop codons. In addition, a 3' primer annealing site is provided so that CDNA synthesis can be accomplished on the mRNA recovered from partitioned
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ligand isolation; systemic polypeptide evolution by reverse translation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2527 tegecgagecegtettegteacegagecegtettegteacegtettegtegeegagecea 2586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence containing a 120 repeat of ACG flanked by fixed fragments
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2347 gcgccgtccggatcgccgtctccgtcggtgagtccgtcggcttcgccgagcctgagcccg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2407 agcccgagcccgagcagctcgccatcgccgtcgccgtcgccgagctcgagtccatcctcg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             tegecgteteegtegecgteaceategeegagteegteteggteteegteaceateggeg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 14; Length 390;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 390 BP; 125 A; 126 C; 133 G; 6 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3.5%; Score 100.8; DB 14;
llarity 61.4%; Pred. No. 5.3e-10;
Conservative 0; Mismatches 102;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               See also AAQ36845-63.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2587 acgccgtcgtcgtcgccggtgtcg 2610
Example 1; Page 84; 98pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             130 TCGTCGTCGTCGTCGTCGTCG 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92US-0829461.
90US-0561968.
91US-0739055.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAF76910 standard; DNA; 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98US-0197649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
Matches 162; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  29-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US6194550-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               23-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    02-AUG-1990;
01-AUG-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            27-FEB-2001
```

```
metabolic pathway; promoter; termination sequence; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0139899.
99US-0140353.
99US-0140354.
99US-0140695.
                                                                                                                                                                                                                                                                                                                      99US-0132484.
                                                                                                                                                                                                                                                                                                                                                                               99US-0134256.
99US-0134218.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0135124.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0139453.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0139461.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99US-0139750.
99US-0139763.
                                                                                                                                                                                                                                                    99US-0130449
                                                                                                                                                                                                                                                                                      99US-0131449
                                                                                                                                                                                                                                                                                                99US-0132048
                                                                                                                                                                                                                                                                                                                                                                                                                                         99US-0134768
                                                                                                                                                                                                                                                                                                                                                                                                                                                    99US-0134941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0135629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99US-0136021
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0136392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0137222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99US-0137528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0137502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99US-0138540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0139119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99US-0139452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99US-0139455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0139456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99US-0139458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0139459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99US-0139460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0139463
                                                                                              25-FEB-2000; 2000EP-0301439
                                                                                                                                                     99US-0125788
99US-0126264
                                                                                                                                                                                                                                                                                                              99US-0132407
                                                                                                                                                                                                                                                                                                                                                                     99US-0132863
                                                                                                                                                                                                                                                                                                                                                                                                                   99US-0134221
                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0134370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0135353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99US-0136782
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0137724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0138847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0139457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99US-0139817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99US-0138094
                           Arabidopsis thaliana.
                                                                                                                                                                                                                                                                                                                     04-MAY-1999;
05-MAY-1999;
06-MAY-1999;
06-MAY-1999;
07-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                14-MAY-1999;
14-MAY-1999;
18-MAY-1999;
19-MAY-1999;
                                                                                                                                          09-MAR-1999;
23-MAR-1999;
25-MAR-1999;
29-MAR-1999;
01-APR-1999;
06-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                               11-MAY-1999;
14-MAY-1999;
14-MAY-1999;
                                                EP1033405-A2
                                                                                                                                                                                                                                                 21-APR-1999;
23-APR-1999;
23-APR-1999;
28-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         07-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       14-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               16-JUN-1999
                                                                       06-SEP-2000
                                                                                                                                                                                                               08-APR-1999
16-APR-1999
                                                                                                                                                                                                                                                                                               30-APR-1999
30-APR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20-MAY-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          21-MAY-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  03-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              04-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      18-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               18-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               24-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                18-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      18-JUN-1999
   The present sequence was used in an example illustrating an invention relating to a method for isolating a polypeptide ligand for a desired target molecule. The method involves synthesising a nucleic acid mixture comprising mRNA having translatable and non-translatable regions and a mixture of nucleic acid.polypeptide copolymers, each comprising the mRNA and a polypeptide encoded by its associated mRNA. The copolymers are partitioned relative to their affinity to the target. The method is termed systemic polypeptide evolution by reverse translation (SPERT). The polypeptides ligands of small molecule targets are useful in assay methods, diagnostic procedures, cell sorting, as inhibitors of target molecule function, as probes, as fury delivery vehicles and modifiers of hormone action and have therapeutic uses as sequestering agents. The target molecules include natural and synthetic polymers, including target molecules such as drugs, metabolites, cofactors and toxins. Polypeptide ignands are isolated and rapidly identified by this method.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                      Isolating a polypeptide ligand to a target molecule, useful for diagnostic assays, comprises partitioning candidate mixtures comprised of ribosome complexes or mRNA.polypeptide copolymers relative to their affinity to the target molecule
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2347 gegecgtecggategecgtetecgteggtgagtecgteggettegeegageetgggeeeg 2406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        tcgccgtctccqtcgccqtcaccatcgccgaqtccgtctcggtctccgtcaccatcggcg 2526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       tcgccgagcccgtcttcgtcaccgagcccgtcttcgtcaccgtcttcgtcgccgagccca 2586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2407 agcccgagcccgagcagctcgccatcgccgtcgccgtcgccgagctcgagtccatcctcg 2466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hybridisation assay; genetic mapping; gene expression control; protein identification; signal transduction pathway;
                                                                                                                                                                                                                                                                                                                                                                                                                              DB 22; Length 390;
                                                                                                                                                                                                                                                                                                                                                                                                                            Score 100.8; DB 22; Length 3 Pred. No. 5.3e-10; 0; Mismatches 102; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Arabidopsis thaliana DNA fragment SEQ ID NO: 55894.
                                                                                                                                                                                                                                                                                                                                                                                         Sequence 390 BP; 125 A; 126 C; 133 G; 6 T; 0 other;
                                                            Smith JD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Example; Column 39; 35pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAC47999 standard; DNA; 784 BP.
                                                           Tuerk C, Pribnow D,
                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 3.5%;
Best Local Similarity 61.4%;
Matches 162; Conservative 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18-OCT-2000 (first entry)
                                                                                 WPI; 2001-243412/25.
GOLD L.
TUERK C.
PRIBNOW D.
SMITH J D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAC47999;
 (GOLD/) C
(TUER/) T
(PRIB/) E
(SMIT/) $
                                                        Gold L,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAC47999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              K K X E X D X C
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
PR 28-JUN-1999; 99US-0140823.

PR 37-JUN-1999; 99US-0140821.

PR 01-JUL-1999; 99US-0141284.

PR 01-JUL-1999; 99US-0141284.

PR 02-JUL-1999; 99US-0141284.

PR 13-JUL-1999; 99US-014286.

PR 13-JUL-1999; 99US-014280.

PR 13-JUL-1999; 99US-014280.

PR 13-JUL-1999; 99US-014381.

PR 13-JUL-1999; 99US-014382.

PR 22-JUL-1999; 99US-014332.

PR 22-JUL-1999; 99US-014433.

PR 22-JUL-1999; 99US-014434.

PR 22-JUL-1999; 99US-014762.

PR 22-JUL-1999; 99US-01476.

PR 22-JUL-1999; 99US-01
```

```
2420 geagetegecategecgtegeegtegecgagetegagtecatectegtegeegteteegt 2479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hybridisation assay; genetic mapping; gene expression control;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 784;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Arabidopsis thaliana DNA fragment SEQ ID NO: 70586.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
3.1%; Score 89; DB 21; L.
Best Local Similarity 72.0%; Pred. No. 8.4e-08;
Matches 116; Conservative 0; Mismatches 45;
990S-0153070
990S-0154018
990S-0154039
990S-0154039
990S-0154039
990S-0155139
990S-0155186
990S-015656
990S-015656
990S-015656
990S-015656
990S-015656
                                                                                                                                                                                                                                                                                                                                                                                                           990S-0161406.
990S-0161359.
990S-0161360.
990S-0161361.
990S-0161920.
                                                                                                                                                                                 99US-0159293.
99US-0159294.
99US-0159295.
                                                                                                                                                                                                                 99US-0159329.
99US-0159330.
99US-0159331.
99US-0159637.
                                                                                                                                                                                                                                                              99US-0159638.
99US-0159584.
99US-0160741.
                                                                                                                                                                                                                                                                                              99US-0160767.
99US-0160768.
99US-0160770.
                                                                                                                                                                                                                                                                                                                                99US-0160814.
99US-0160815.
99US-0160980.
                                                                                                                                                                                                                                                                                                                                                                           99US-0160989
99US-0161404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99US-0161993
99US-0162142
                                                                                                                                                                                                                                                                                                                                                                                                  99US-0161405
                                                                                                                                                                                                                                                                                                                                                                 99US-0160981
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18-OCT-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAC52204/c
ID AAC52204 standard;
10.5EP-1999;
13.5EP-1999;
16.5EP-1999;
20.5EP-1999;
22.5EP-1999;
23.5EP-1999;
24.5EP-1999;
24.5EP-1999;
26.0CT-1999;
66.0CT-1999;
66.0CT-1999;
67.0CT-1999;
68.0CT-1999;
                                                                                                                                                                                                                                                 14 OCT - 1999;
14 OCT - 1999;
18 OCT - 1999;
21 OCT - 1999;
21 OCT - 1999;
21 OCT - 1999;
21 OCT - 1999;
22 OCT - 1999;
22 OCT - 1999;
                                                                                                                                                                                                                            14-OCT-1999;
14-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                 25-OCT-1999;
25-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAC52204;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
```

us-09-917-376-2.rng

																		•																				
	-																		-	-														-				
,																																						
Arabidopsis thaliana.		2000EP-0301439.	99US-0121825. 99US-0123180.	990S-0125788.	99US-0126264.	99US-0127462.	99US-0128714.	99US-0129845.	99US-0130449.	990S-0130810.	99US-0131449.	99US-0132407.	99US-0132484.	99US-0132486.	990S-0132487. 990S-0132863.	99US-0134256.	99US-0134219.	99US-0134221. 99US-0134370.	99US-0134768.	99US-0135124.	99US-0135353. 99US-0135629.	99US-0136021.	99US-0136782.	99US-0137222. 99US-0137528.	99US-0137502.	99US-0138094.	99US-0138540. 99US-0138847.	99US-0139119.	990S-0139453.	990S-0139492.	99US-0139455.	99US-0139457.	99US-0139458.	99US-0139460.	99US-0139461.	99US-0139463.	99US-0139750. 99US-0139763.	99US-0139817.
	06-SEP-2000.	FEB-2000; 20	; 660; 660;	, ,	. 66	660	99;	666:	66	900	. 66	999;	660 600	,66	. 66 . 66	660			66)	,66		. 66	,66		660	99;	99;	.66	, 66	. 66	. 660	666	166	99;	. 66	666	, 66 96;	:66

PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 20-JUN-1999; 99US-0140823.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 02-JUL-1999; 99US-0142065.
PR 03-JUL-1999; 99US-0142065.
PR 13-JUL-1999; 99US-0142020.
PR 13-JUL-1999; 99US-0142020.
PR 13-JUL-1999; 99US-0142020.
PR 13-JUL-1999; 99US-0144332.
PR 13-JUL-1999; 99US-0144332.
PR 13-JUL-1999; 99US-0144332.
PR 13-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144333.
PR 20-JUL-1999; 99US-0144333.
PR 20-JUL-1999; 99US-0144334.
PR 21-JUL-1999; 99US-0144334.
PR 22-JUL-1999; 99US-014508.
PR 22-JUL-1999; 99US-014508.
PR 23-JUL-1999; 99US-014508.
PR 23-JUL-1999; 99US-014508.
PR 23-JUL-1999; 99US-014508.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145149.
PR 23-JUL-1999; 99US-01499.
PR 23

ဖ

```
Sequences AASJOON TEPPESSENT DAY MAN INCIDENCE.

Sequences AASJOON TEPPESSENT DAY MAN INCIDENCE.

Associated DNA sequences are used in the treatment, prevention and diagnosis of medical conditions caused by P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and consecrated by P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be conseduce expression and activity of P. acnes polypeptides and therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by consymptides shown in AAU44158-AAU4458 and AAU67488-AAU67490.

Note: The sequence data for this partent did not form part of the printed specification, but was obtained in electronic format directly from WIPO
SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis; uveitis; endophthalmitis; bone; joint; central nervous system; ELISA; inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay; dermatological; osteopathic; neuroprotectant; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            14100 CCGGCTCCGACACCGACACCGACACCGGCTCCGACACCGACACCCCACT 34041
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          34220 CCGACTCCTACGCCGACCCCGACTCCTACGCCGACCCCGACTCCTACGCCGACCCCGACT 34161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     34160 CCTACGCCGACACCGACACCGACACCGACACCGACACCGGCTCCGACACCGGCTCCGACA 34101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2440 cogtogocgagetegagtecatectegtegeegteteegtegeegteaceategeegagt 2499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2500 ccgtctcggtctccgtcaccatcggcgtcgccgagcccgtcttcgtcaccgagcccgtct 2559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2380 ccgtcggcttcgccgagcctgagcccgagcccgagcccgagcagctcgccatcgccgtcg 2439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Propionibacterium acnes polypeptides and nucleic acids useful for vaccinating against and diagnosing infections, especially useful for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 37286;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 37286 BP; 7034 A; 10065 C; 11922 G; 8264 T; 1 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequences AAS59506-AAS59804 represent DNA molecules encoding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mitcham JL, Wang SS, Bhatia A;
Jen S, Carter D;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2560 tcgtcaccgtcttcgtcgccgagcccaacgccgtcgtcgcc 2603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 3.0%; Score 86.4; DB 23; Best Local Similarity 61.6%; Pred. No. 2.7e-07; Matches 138; Conservative 0; Mismatches 86;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 1; SEQ ID No 17; 1069pp; English.
                                                                                                                                                                                                                                                                                                                                                             21-APR-2000; 2000US-199047P.
02-JUN-2000; 2000US-208841P.
07-JUL-2000; 2000US-216747P.
                                                                                                                                                                                                                                                                                                            20-APR-2001; 2001WO-US12865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Skeiky YAW, Persing DH, M
L'maisonneuve J, 'Zhang Y,
                                                                                                                                          Propionibacterium acnes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             treating acne vulgaris
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2001-616774/71.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (CORI-) CORIXA CORP
                                                                                                                                                                                              WO200181581-A2
                                                                                                                                                                                                                                                     01-NOV-2001
     NAMES OF THE PROPERTY OF THE P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         gcagctcgccatcgccgtcgccgtcgccgagctcgagtccatcctcgtcgccgtctcgt 2479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1153 GCTCTCTTCGTCTCCGTCTCCGTCTCCGTCTCCGTCTCCGTCTCCGTCTCCGTTTCCGT 1094
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Propionibacterium acnes immunogenic protein encoding DNA #17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 2011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       45; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2540 cttcgtcaccgagcccgtcttcgtcaccgtcttcgtcgccg 2580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1033 CTCCGTCTCCGTCTCCGTCTCCATCTCTCTCTCCG 993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 89; DB 21;
Pred. No. 8.5e-08;
0; Mismatches 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAS59522/c
ID AAS59522 standard; DNA; 37286 BP
                         990S-0153070
990S-0153070
990S-0154018
990S-0154039
990S-0155139
990S-0155186
990S-0155486
990S-015656
990S-015656
990S-0156596
990S-0157117
990S-0157117
990S-015929
990S-015929
990S-015929
990S-015939
990S-015938
990S-015938
990S-015938
990S-015938
990S-015938
990S-015938
990S-015938
990S-015938
990S-015938
990S-0160768
990S-0160815
990S-0160815
990S-0160815
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     990s-0161359.
990s-0161360.
990s-0161361.
990s-0161920.
990s-0161992.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99US-0162142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches 116; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                              08-0CT-1999;
12-0CT-1999;
13-0CT-1999;
13-0CT-1999;
13-0CT-1999;
14-0CT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                14-0CT-1999;
14-0CT-1999;
14-0CT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21-OCT-1999;
22-OCT-1999;
22-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2-OCT-1999;
5-OCT-1999;
5-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   13-FEB-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1-OCT-1999;
1-OCT-1999;
                                                                                                                                                                                                                                                                                                      04-0CT-1999;
05-0CT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   14-OCT-1999;
18-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                06-0CT-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21-OCT-1999
21-OCT-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26-OCT-1999
26-OCT-1999
                                                                                                                                                                                                                                                                                                                                                                                          -OCT-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             26-0CT-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              25-OCT-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAS59522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21-OCT-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
```

ö

XEXEXEX

ò a q ŏ

ò

```
99US-0139458
99US-0139469
99US-0139461
99US-0139461
99US-0139462
99US-0139750
99US-0139750
99US-01391763
                                                                                                                                                               990S - 0140353
990S - 0140354
990S - 0140835
990S - 0140891
990S - 0141287
990S - 0141287
990S - 0142055
990S - 0142059
990S - 0142059
990S - 0142920
990S - 0142920
990S - 014394
990S - 014394
                                                                                                                                                                                                                                                                                                                                                                                    990S-0144085.
990S-0144086.
990S-0144325.
990S-0144331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9903-0144884
9903-0145886
9903-0145086
9903-0145085
9903-0145087
9903-0145089
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               990S-0145145.
990S-0145218.
990S-0145224.
990S-0145276.
990S-0145913.
990S-0145913.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9905-0145951
9905-0146386
9905-0146389
9905-0146389
9905-0147038
9905-0147304
9905-0147302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0144335.
99US-0144352.
99US-0144632.
  99US-0139456
99US-0139457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99US-0147935
99US-0148171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0149368
99US-0149175
                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0144333
99US-0144334
18 - JUN - 1999;

22 - JUN - 1999;

23 - JUN - 1999;

24 - JUN - 1999;

25 - JUN - 1999;

26 - JUN - 1999;

27 - JUN - 1999;

28 - JUN - 1999;

29 - JUN - 1999;

20 - JUL - 1999;

21 - JUL - 1999;

22 - JUL - 1999;

23 - JUL - 1999;

24 - JUL - 1999;

25 - JUL - 1999;

26 - JUL - 1999;

27 - JUL - 1999;

28 - JUL - 1999;

29 - JUL - 1999;

20 - JUL - 1999;

20 - JUL - 1999;

21 - JUL - 1999;

22 - JUL - 1999;

23 - JUL - 1999;

24 - JUL - 1999;

25 - JUL - 1999;

26 - JUL - 1999;

27 - JUL - 1999;

28 - JUL - 1999;

29 - JUL - 1999;

20 - JUL - 1999;

20 - JUL - 1999;

20 - JUL - 1999;

21 - JUL - 1999;

21 - JUL - 1999;

21 - JUL - 1999;

22 - JUL - 1999;

23 - JUL - 1999;

24 - JUL - 1999;

25 - JUL - 1999;

26 - JUL - 1999;

27 - JUL - 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22 - 70L - 1999;
22 - 70L - 1999;
23 - 70L - 1999;
23 - 70L - 1999;
23 - 70L - 1999;
26 - 70L - 1999;
27 - 70L - 1999;
27 - 70L - 1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-70L-1999;
19-70L-1999;
19-70L-1999;
20-70L-1999;
20-70L-1999;
21-70L-1999;
21-70L-1999;
21-70L-1999;
21-70L-1999;
21-70L-1999;
21-70L-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     02-AUG-1999;
03-AUG-1999;
04-AUG-1999;
04-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28-JUL-1999;
02-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          02-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           05-AUG-1999;
05-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      06-AUG-1999;
06-AUG-1999;
  Hybridisation assay, genetic mapping, gene expression control, protein identification, signal transduction pathway; metabolic pathway; promoter; termination sequence; ss.
Db 34040 CCGACACCGACACCCAACGCCCACCCGACCCCCAACCCC 33997
                                                                                                                                                  Arabidopsis thaliana DNA fragment SEQ ID NO: 42845.
                                                                  AAC4444 standard; DNA; 777 BP
                                                                                                                                                                                                                                                                                                                                          990S-0121825.
990S-0123180.
990S-0125788.
990S-0125788.
990S-0126765.
990S-0126765.
990S-0128714.
990S-0130877.
990S-0130891.
990S-0130891.
990S-0131448.
990S-0132486.
990S-0132486.
990S-0132486.
990S-0132486.
990S-013486.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                990S-0137724
990S-0138094
990S-0138847
990S-0138847
990S-0139452
990S-0139452
990S-0139452
                                                                                                                                                                                                                                                                                                                 25-FEB-2000; 2000EP-0301439
                                                                                                                       18-OCT-2000 (first entry)
                                                                                                                                                                                                                                  Arabidopsis thaliana
                                                                                                                                                                                                                                                           EP1033405-A2
                                                                                                                                                                                                                                                                                                                                                                                                25-MAR-1999

29-MAR-1999

06-APR-1999

10-APR-1999

21-APR-1999

21-APR-1999

23-APR-1999

23-APR-1999

23-APR-1999

30-APR-1999

06-MAY-1999

06-MAY-1999

06-MAY-1999

11-MAY-1999

11-MAY-1999

11-MAY-1999

11-MAY-1999

14-MAY-1999

14-MAY-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24 - MAY - 1999;
25 - MAY - 1999;
28 - MAY - 1999;
01 - JUN - 1999;
04 - JUN - 1999;
04 - JUN - 1999;
08 - JUN - 1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10-JUN-1999;
10-JUN-1999;
14-JUN-1999;
16-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18-JUN-1999;
18-JUN-1999;
                                                                                                                                                                                                                                                                                      06-SEP-2000
                                                                                             AAC4444;
                                       7
                                                      AAC44444
                                         RESULT
                                                                                 2
```

```
Hybridisation assay; genetic mapping; gene expression control; protein identification; signal transduction pathway; metabolic pathway; promoter; termination sequence; ss.
                                                                                                                                                                                       Arabidopsis thaliana DNA fragment SEQ ID NO: 44942.
   172 ctccgtctccgtctccatctctctctccggt 202
                                                                                                                                                                                                                                                                                                                                                                                                      990S-0121825
990S-0123180
990S-0125788
990S-0125788
990S-0126264
990S-0128234
990S-0128234
990S-0128945
990S-0139845
990S-0130077
990S-0130891
990S-0130891
990S-0132486
990S-0132486
990S-0132486
990S-0132486
990S-0132486
990S-0132486
990S-0132486
990S-0132486
990S-0132486
990S-0132487
990S-0132487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99US-0135629.
99US-0136021.
99US-0136392.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99US-0137222.
99US-0137528.
99US-0137502.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99US-0137724.
99US-0138094.
99US-0138540.
99US-0138847.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0139452.
99US-0139453.
99US-0139492.
99US-0139454.
                                                                                             AAC45005 standard; DNA; 796
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0134768
99US-0134941
                                                                                                                                                                                                                                                                                                                                                                           2000EP-0301439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0136782
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99US-0139119
                                                                                                                                                                                                                                                                               Arabidopsis thaliana
                                                                                                                                                                                                                                                                                                                                                                         25-FEB-2000;
                                                                                                                                                                                                                                                                                                                                                                                                    25 - FEB - 1999

05 - MAR - 1999

23 - MAR - 1999

25 - MAR - 1999

01 - APR - 1999

01 - APR - 1999

06 - APR - 1999

19 - APR - 1999

10 - APR - 1999

10 - APR - 1999

11 - APR - 1999

12 - APR - 1999

13 - APR - 1999

14 - MAY - 1999

15 - MAY - 1999

16 - MAY - 1999

17 - MAY - 1999

18 - MAY - 1999

19 - MAY - 1999

19 - MAY - 1999

20 - MAY - 1999

21 - MAY - 1999

22 - MAY - 1999

23 - MAY - 1999

24 - MAY - 1999

25 - MAY - 1999

27 - MAY - 1999

28 - MAY - 1999

28 - MAY - 1999

27 - MAY - 1999

28 - MAY - 1999
                                                                                                                                                                                                                                                                                                              EP1033405-A2
                                                                                                                                                        18-OCT-2000
                                                                                                                                                                                                                                                                                                                                            06-SEP-2000
                                                                                                                           AAC45005;
                                                                œ
                                                                              AAC45005
                                                                RESULT
                   С
                                                                                              ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2456 giccatccicgtcgccgiccgicgccgicaccaicgccgagiccggictccggi 2515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            52 gectetettegteteegteteegteteegteteegteteegteteegteteegteteegt 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 85.4; DB 21; Length 777; Pred. No. 3.9e-07;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      41;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2576 cgccgagcccaacgccgtcgtcgccggt 2606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3.0%;
72.8%;
990S-0149426.
990S-0149722.
990S-0149723.
990S-0149923.
990S-0149930.
990S-0150566.
990S-015066.
990S-0151066.
990S-0151066.
990S-0151069.
990S-0151089.
                                                                                                                                                                                                                  990S-0152363.
990S-0153070.
990S-0154018.
990S-0154039.
990S-0154779.
990S-0155486.
990S-0155486.
990S-0155486.
990S-0155486.
990S-0156759.
990S-0156759.
990S-01598629.
990S-01598629.
990S-01598629.
990S-01598629.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0159330.
99US-0159637.
99US-0159637.
99US-0159584.
99US-0160741.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     990S-0160768.
990S-0160770.
990S-0160814.
990S-0160980.
990S-0160980.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0161405.
99US-0161406.
99US-0161359.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0161360.
99US-0161361.
99US-0161920.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0161992.
99US-0161993.
99US-0162142.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0161404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 72.8
Matches 110; Conservative
20-AuG-1999;
20-AuG-1999;
20-AuG-1999;
21-AuG-1999;
22-AuG-1999;
22-AuG-1999;
22-AuG-1999;
23-AuG-1999;
22-AuG-1999;
23-AuG-1999;
23-AuG-1999;
23-AuG-1999;
23-AuG-1999;
23-AuG-1999;
24-AuG-1999;
25-AuG-1999;
26-AuG-1999;
26-Au
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
```

```
990S - 0139455
990S - 0139456
990S - 0139456
990S - 0139459
990S - 0139462
990S - 0139463
990S - 0140354
990S - 0140354
990S - 0141287
990S - 014287
990S - 014287
990S - 0144335
990S - 0144335
990S - 0144335
990S - 0144336
990S - 0145086
990S - 0145086
990S - 0145086
990S - 0145189
990S - 0147192
990S - 0147192
990S - 0147192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99US-0148565.
99US-0148684.
99US-0149368.
                                                18 - 70N - 1999

22 - 70N - 1999

23 - 70N - 1999

24 - 70N - 1999

26 - 70N - 1999

27 - 70N - 1999

28 - 70N - 1999

29 - 70N - 1999

20 - 70N - 1999

20 - 70N - 1999

20 - 70N - 1999

21 - 70L - 1999

22 - 70L - 1999

23 - 70L - 1999

24 - 70L - 1999

25 - 70L - 1999

26 - 70L - 1999

27 - 70L - 1999

28 - 70L - 1999

29 - 70L - 1999

20 - 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22 - JUL - 1999;
23 - JUL - 1999;
23 - JUL - 1999;
23 - JUL - 1999;
26 - JUL - 1999;
27 - JUL - 1999;
27 - JUL - 1999;
28 - JUL - 1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             04 - AUG - 1999

04 - AUG - 1999

05 - AUG - 1999

06 - AUG - 1999

06 - AUG - 1999

09 - AUG - 1999

10 - AUG - 1999

10 - AUG - 1999

10 - AUG - 1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11-AGG-1999;
12-AGG-1999;
13-AGG-1999;
13-AGG-1999;
16-AGG-1999;
```

Query Match 3.0%; Score 85; DB 21; Length 796; Best Local Similarity 71.3%; Pred. No. 4.7e-07; Matches 112; Conservative 0; Mismatches 45; Indels 9905 - 0149426 9908 - 0149429 9908 - 0149923 9908 - 0149923 9908 - 0149923 9908 - 015066 9908 - 0151066 9908 - 0151066 9908 - 0151080 9908 - 0151080 9908 - 0151080 9908 - 0151080 9908 - 0151080 9908 - 0151080 9908 - 0155659 9908 - 0156659 9908 - 0156638 9908 - 0156638 9908 - 0156638 9908 - 015923 9908 - 015923 9908 - 015923 9908 - 015923 9908 - 015923 9908 - 015923 9908 - 015923 9908 - 015923 9908 - 015923 9908 - 015923 9908 - 015923 9908 - 015923 9908 - 015923 9908 - 015923 9908 - 015923 9908 - 016074 9908 - 016076 9908 - 016076 9908 - 016076 9908 - 016098 9908 - 01601406 9908 - 0160136 9908 - 016136 9908 - 016136 9908 - 016136 9908 - 016136 17-AUG-1999;
20-AUG-1999;
20-AUG-1999;
20-AUG-1999;
23-AUG-1999;
23-AUG-1999;
23-AUG-1999;
25-AUG-1999;
27-AUG-1999;
28-SEP-1999;
28-SEP-1999;
29-SEP-1999;
20-CT-1999;
20-CT-1999;
21-CT-1999;
21-CT-19 21-0CT-1999; 21-0CT-1999; 21-0CT-1999; 21-0CT-1999; 22-0CT-1999; 22-0CT-1999; 22-0CT-1999; 25-0CT-1999; 25-0CT-1999; 26-0CT-1999; 26-0CT-1999; 26-0CT-1999; 28-0CT-1999; 28-0CT-1999; 2424 Db

ö

Gaps

gtcaccatcgccgagtccgtctcggtctccgtcaccatcggcgtcgccgagcccgtcttc 2543

51 2484 111

> Qγ g

170

ò g

```
4194 gcggtatgtgactgggcacagataggggcaagcaatgtgacattcaattttgtgaagctg 4253
   2713 ttgtcgacggtgacggtgcggtactggttcacccgggatggtggctcgtcgacactggtg
                                                                                                ccgacagtgacggttactgtgactccgacaccgacaccaacagccgacgacgacac
                                                                                                                                 2596 tegtegeeggtgtegg---gtgggggtgaaggtgeagtataagaataatgatteggegeeg
                                                                                                                                                                                                                  ggtgataatcagatcaagccgggtttgcaggtggtgaataccgggtcgtcgtcggtggat
                                                          ccgtcttcgtcaccgagcccgtcttcgtcaccgtcttcgtcgccgagcccaacgccgtcg
                                                                                                                                                                      gogacacctggcacgggaagtggtttgaaggtactatacaagaacaatgagacaagtgcg
                                                                                                                                                                                                                                                                                                                                                                 tacaactgtgactgggcggcgatcgggtgtgggaatatccgcgcctcgttcggctcggtg
                                                                                                                                                                                                                                                                                                                                                                                                                                            2833 aacccggcgacgccgacggcggacacctacctgcag 2868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4254 agcagcggagtgagtggagcggattattacttggag 4289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAD26525 standard; DNA; 11707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98US-0136574
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /*tag= b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Anderson
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Farrington GK, Anderso
Morgan H, Williams DP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2002-081780/11.
P-PSDB; AAE16323.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19-SEP-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JS6294366-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAD26525;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10
                     3897
                                                          2536
                                                                                                3957
                                                                                                                                                                        4017
                                                                                                                                                                                                                2653
                                                                                                                                                                                                                                                                                                                                                                   2773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAD26525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                          ğ
                                                                                                                                   δ
                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The invention relates to a recombinant cellulase active protein free of proteinases of native thermophilic and alkalinophilic origin, comprising the truncated sequences Cel B5, Cel B45, Cel E1, Cel E1/2, Cel 1/2/3, Cel 6 or Cel B3/B5, or a stability region from one of the defined full-length sequences, or functional equivalents. Cel B5 extends from amino acid K635 to N1426 in the sequence shown in AAY13492; Cel E1 extends from amino acid K635 to N1426 in the sequence shown in AAY13492; Cel E1 extends from amino acid X39 to D481, Cel E1/2 extends from x39 to G635, Cel E1/2/3 extends from amino acid V1233 to K1751 and the stability region extends from amino acid V1233 to K1751 and the stability region extends from amino acid E482 to G635 in the sequence shown in AAY13493. Cel E3/B5 is shown in AAY13494. The new conceptually backstaining or graying, for use on cellulosic materials including cotton-containing fabrics. They are especially useful for preventing redeposition of colorant during stonewashing, and for processing of textiles where cellulose breakdown is required. The new truncated enzymes show reduced redeposition of dye compared to using
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ;
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3837 cogacacoggicacogacocogacatogacgicogacacocacacotacaacocacotaca 3896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            cogtogcogtoaccatcgccgagtccgtctcggtctccgtcaccatcggcgtcgccgagc 2535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     cogagoagotogocatogocogtogocogagotogagtocatoctogtogocogtot 2475
                                                                                                                                                                                                                                                                                        El; Cel El/2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                    Cellulase; proteinase; truncated; Cel B5; Cel B4/5; Cel E1; Cel Cel 1/2/3; Cel 6; Cel E3/B5; laundry detergent; stain-removing; cotton-containing fabric; stonewashing; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9:
                                                                                                                                                                                                                                                   DNA sequence encoding truncated cellulases Cel B4/5 and Cel B5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         proteins, useful in detergents and for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 11706;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 11706 BP; 3828 A; 1994 C; 2994 G; 2890 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 80; DB 20; Length 11
Pred. No. 4.1e-06;
0; Mismatches 215; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Farrington GK;
t PL, Daniels RM,
Williams DP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Disclosure; Page 20-23; 65pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              non-truncated cellulase compositions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CLRN ) CLARIANT FINANCE BVI LTD.
                                                                                                                                   AAX55661 standard; DNA; 11706 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             denim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98EP-0810919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  97US-0932571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bergquist PL,
                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New truncated cellulase producing 'stonewashed'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Morgan H,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          al Similarity
235; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1999-315403/27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P-PSDB; AAY13492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               15-SEP-1998;
                                                                                                                                                                                                              30-JUL-1999
                                                                                                                                                                                                                                                                                                                                                                 Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .9-SEP-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Anderson P,
                                                                                                                                                                                                                                                                                                                                                                                                      EP921188-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                          09-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gibbs MD,
                                                                                                                                                                        AAX55661;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2476
                                                                                                                 AAX55661
                                                                                                RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
```

BP.

```
Active cellulase protein; alkalinophilic; textile processing; proteinase; detergent additive; stonewashed appearance; cotton-containing denim; CelB5; thermophilic; commercial detergent; celB gene; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New cellulase active protein, useful in textile processing or commercial detergents, e.g. for improving the feel or appearance of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gibbs MD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Daniels R,
                                                                                                                                                                                                                                                                                                                                                                                            /product= "DNA encoding B4/5 protein"
9255.10526
/*tag= c
/product= "DNA encoding B/5 protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             P, Bergquist P,
                                                                                                                                                                                                                                                                                                                   protein"
Active cellulase protein, celB gene.
                                                                                                                                                                                                                             Location/Qualifiers
6255..10535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (CLRN ) CLARIANT FINANCE BVI LTD.
                                                                                                                                                                                                                                                                                                          /product= "CelB
8601..10532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               97US-0932571.
```

ò

ð

Evaluating strain variation of Mycobacterium tuberculosis, comprises determining the nucleotide sequence of the strain at positions in the genome corresponding to positions where M. tuberculosis strains CDC

1551 and H37Rv differ

Venter JC;

Fraser CM,

Fleischmann RD, White OR,

WPI; 2001-647261/74

GENOMIC RES

(GENO-) INST

98US-0103840 98US-0103840

24 - JUN - 1998; 24-JUN-1998;

US6294328-B1

25-SEP-2001

```
2;
                                                                                                                                                        the
                                                                                                     The present invention relates to a cellulase active protein, which is substantially free of proteinases of native thermophilic and alkalinophilic origin, where the cellulase active protein consists of the CelBS amino acid sequence. The cellulase active protein is useful for treating cellulosic materials including cotton-containing fabrics, as detergent additives. The cellulase active protein is also useful for improving the feel and/or appearance of cotton-containing fabrics, for removing surface fibers from cotton-containing whits or for imparting stonewashed appearance to cotton-containing denims. The present proteins are stable under condition of alkaline pH and elevated temperatures, thus suitable for textile processing and in commercial detergents.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4077 agcacaagttctataaggccgtggtttaagatagtgaatggaggcagcagcagtgttgat 4136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3837 cogacacoggoacogacocogacatogacgocogacacoaacacotacaacoacoacotgoa 3896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4137 cttagcagggttaagataagatactggtacacagtggatggtgac---aagccacagagt 4193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4194 gcggtatgtgactgggcacagataggggcaaggcaatgtgacattcaattttgtgaagctg 4253
      ЬH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome; variation; epidemiology; patient treatment; epidemic monitoring; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2416 ccgagcagctcgccatcgccgtcgccgtcgccgagctcgagtccatcctcgtcgccgtct
cotton-containing fabrics, is stable under conditions of alkaline and elevated temperatures -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3897 ccgacatcagccccgacaccgagcccaacagtgacagcaacgccgactccaacgccgacg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2536 ccgtcttcgtcaccgagcccgtcttcgtcaccgtcttcgtcgccgagcccaacgccgtcg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2476 ccgtcgccgtcaccatcgccgagtccgtctcggtctccgtcaccatcggcgtcgccgagc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             tcgtcgccggtgtcgg---gtggggtgaaggtgcagtataagaataatgattcggcgccg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4017 gggacacctggcacgggaagtggtttgaaggtactatacaagaacaatgagacaagtgcg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ggtgataatcagatcaagccgggtttgcaggtggtgaataccgggtcgtcgtcggtggat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2713 ttgtcgacggtgacggtgcggtactggttcacccgggatggtggctcgtcgacactggtg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2773 tacaactgtgactgggcggcggtcggtgtggggatatccgcgcctcgttcggctcggtg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 11707;
                                                                                                                                                                                                                                                                                                                                                                                               Seguence 11707 BP; 3827 A; 1995 C; 2994 G; 2891 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 80; DB 24; Length 11
Pred. No. 4.1e-06;
0; Mismatches 215; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2833 aacccggcgacgccgacggcggacacctacctgcag 2868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4254 agcagcggagtgagtggagcggattattacttggag 4289
                                                              Disclosure; Column 27-38; 61pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BP
                                                                                                                                                                                                                                                                                                                                                       The present sequence is celB gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAI99683 standard; DNA; 4403765
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 2.8%;
Best Local Similarity 51.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mycobacterium tuberculosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             235; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 15-JAN-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAI99683;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAI99683/c
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
3928646 GCCCGCCGGTGCCGCCGGTCCCCGTTGATGCCGCCCGGCCCGGCTTGCCG 3928587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Db 3928766 CCGTCGCCGGGAACACCGGCGCCCGAATCCGGCGTTGCCGCCCGGCCCCGCCGATGCCGCCT 3928707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3928586 CCGACCCCCTTGGCCGCCTTGGCCGCCGAAGCCGTTGCCGTCTTGGGAGAGGGCG 3928527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3928466 CCCGCGGTTCCGCCGGTGTTGGCCAGGCCGCCGACGCCGGCGTCGCCGGTCCCGGCG 3928407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           The invention relates to evaluating strain variation within and between different populations of the tuberculosis bacterial pathogen, Mycobacterium tuberculosis or related Mycobacterium by determining the nucleotide sequence of the first strain at positions in the complete sequence of the genome that correspond to positions that differ in the nucleotide sequences of M. tuberculosis strains CDC 1551 (AA199683) and H37Rv (AA199682). The method is useful for evaluating strain variation of M. tuberculosis and has valuable application in the fields of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2344 ggtgcgccgtccggatcgccgtctccgtcgtcggtgagtccgtcggcttcgccgagcctgagc 2403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2404 cogagecegagecegageagetegeeategeegtegeegtegeegtegeegagetegagteeatee 2463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2584 ccaacgccgtcgtcgtcgtcggtgtcgggtggggtgaaggtgcagtataagaataatgat 2643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2524 gcgtcgccgagcccgtcttcgtcaccgagcccgtcttcgtcaccgtcttcgtcgccgagc 2583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2704 tcggtggatttgtcgacggtgacggtgcggtactggttcacccggggatggtggctcgtcg 2763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2764 acactggtgtacaactgtgactgggcggcgatcgggtgtggggaatatccgcgcctcgttc 2823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2644 toggcgccgggtgataatcagatcaagccgggtttgcaggtggtgaataccgggtcgtcg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2464 tcgtcgccgtctccgtcgccgtcaccatcgccgagtccgtctcggtctccgtcaccatcg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 4403765;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 2.8%; Score 79; DB 22; Length 44 Best Local Similarity 47.2%; Pred. No. 6.8e-06; Matches 241; Conservative 0; Mismatches 270; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       monitoring.
δλ
```

```
Sequence 58857 BP; 7256 A; 25139 C; 19353 G; 7109 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BLM gene cluster; bleomycin gene cluster; polyketide metabolite; bleomycin; bleomycin analogue; holo-carrier protein; thiazolidine; thiazoline; bithiazoline; microbial metabolite; sugar; ss.
                                                                                                                                                                                                                                                       Nucleotide sequence of the bleomycin (BLM) gene cluster ORFs 8-30.
                                                                                                                                                                                                                                                                                                                                                                                                                                                /*tag= a
/transl_except= (pos: 1..3, aa: Met)
/note= "ORF 30; encodes AAB07556"
561..2309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /transl_except= (pos: 1..3, aa: Met)
/note= "ORF 29; encodes AAB07557"
2767..3486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /transl_except= (pos: 1..3, aa: Met)
/note= "ORF 27; encodes AAB07559"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /transl_except= (pos: 1.3, aa: Met)
/note= "ORF 23; encodes AAB07563"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /transl_except= (pos: 1..3, aa: Met)
/note= "ORF 18; encodes AAB07568"
99301..47181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /transl_except= (pos: 1..3, aa: Met)
/note= "ORF 20; encodes AAB07566"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /transl_except= (pos: 1..3, aa: Met)
/note= "ORF 19; encodes AAB07567"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12391..15491
/*tag= f
/note= "ORF 25; encodes AAB07561"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note= "ORF 23; encodes AAB07563"
44663..32690
/*tag= i
1000ce= "ORF 22; encodes AAB07564"
7*tag= j
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /*tag= c
/note= "ORF 28; encodes AAB07558"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  *tag= g
/note= "ORF 24; encodes AAB07562"
11010..24666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note- "ORF 17; encodes AAB07569"
47178..49985
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  note= "ORF 21; encodes AAB07565"
14827..35804
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /*tag= e
/note= "ORF 26; encodes AAB07560"
                              Db 3928346 GCCCGCCGTTGCCGCCGTCACCGCTGGCGG 3928316
2824 ggctcggtgaacccggcgacgccggcgg 2854
                                                                                                                                                                                                                                                                                                                                                                                                          Location/Qualifiers
223..564
                                                                                                                                      AAA58471 standard; DNA; 58857 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5488..21013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /note= "ORF 35818..37302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7299..39215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note= "ORF
5806..12294
                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ..5593
                                                                                                                                                                                                                                                                                                                                                                      Streptomyces verticillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 *tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          *tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     *tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 *tag=
                                                                                                                                                                                                                 31-OCT-2000
                                                                                                                                                                           AAA58471;
                                                                                                                  AAA58471/
                                                                                                                                                                                                                                                                                                                                                                                                          Key
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CDS
    ò
                                                                                                                                                                           A CONTRACTOR OF THE CONTRACTOR
```

```
The present sequence represents the BLM (Bleomycin) gene cluster, containing open reading frames (ORFs) 8-30. The proteins encoded by the gene cluster are useful for producing peptides and/or polyketide metabolites, especially bleomycin or bleomycin analogues. They are also useful for chemically modifying biological molecules to produce branched methyl groups, and for coupling amino acids and fatty acids. They may be reacted with an apo-carrier protein and coenzyme A to produce a holo-carrier protein. The BLM gene cluster or catalytic domains can be used individually or collectively to produce thiazoline, thiazoline, bithiazoline and bithiazoline-containing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  microbial metabolites. The BLM gene cluster may also be used to produce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P-PSDB; AABO7556, AABO7557, AABO7558, AABO7559, AABO7560, AABO7561, AABO7562, AABO7563, AABO7564, AABO7565, AABO7566, AABO7569, AABO7570, AABO7571, AABO7572, AABO7573, AABO7574, AABO7575, AABO7576, AABO7577, AABO7578.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     New bleomycin gene cluster components useful for peptide and/or polyketide metabolites, especially bleomycin, production and for chemically modifying biological molecules -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /*tag= v
/transl_except= (pos: 1..3, aa: Met)
/note= "ORF 9; encodes AAB07577"
57583..58857
/transl_except= (pos: 1..3, aa: Met)
/note= "ORF 16; encodes AAB07570"
19982..51001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    except= (pos: 1..3, aa: Met)
"ORF 10; encodes AAB07576"
                                                                                                        /transl_except= (pos: 1..3, aa: Met)
/note= "ORF 15; encodes AAB07571"
                                                                                                                                                                                                                 /transl_except= (pos: 1..3, aa: Met)
/note= "ORF 14; encodes AAB07572"
                                                                                                                                                                                                                                                                                                                            /note= "ORF 13; encodes AAB07573"
53018..54190
                                                                                                                                                                                                                                                                                                                                                                                                           "ORF 12; encodes AAB07574"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "ORF 11; encodes AAB07575"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     note- "ORF 8; encodes AAB07578"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Edwards DJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Claim 8; Page 97-136; 162pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Chen M,
                                                                                                                                       /note= "ORF 1
50998..52386
                                                                                                                                                                                                                                                /note- "ORF 1
52383..52946
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note= "ORF 1
56090..57586
                                                                                                                                                                                                                                                                                                                                                                                                                                       ..55824
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ..56093
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      06-JAN-2000; 2000WO-US00445.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      06-JAN-1999; 99US-0115435.
05-FEB-1999; 99US-0118848.
05-JAN-2000; 2000US-0477962.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sanchez C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (REGC ) UNIV CALIFORNIA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         transl_
                                                                                    *tag=
                                                                                                                                                                                                *tag=
                                                                                                                                                                                                                                                                                                        *tag=
                                                                                                                                                                                                                                                                                                                                                                                    *tag=
                                                                                                                                                                                                                                                                                                                                                                                                              note=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   *tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                *tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             note-
                                                                                                                                                                                                                                                                                                                                                                                                                                       54187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2000-465974/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Du L,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200040704-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Shen B,
                                                        CDS
                                                                                                                                                                 CDS
                                                                                                                                                                                                                                                                         CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CDS
                                                                                                                                                                                                                                                                                                                                                     CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CDS
```

ö

Gaps

ö

2467

79

```
diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Fragments of chemically modified genes associated with tumour suppressor genes and oncogenes, useful in designing primers and probes for
                                                                                                                                                                                                                                                             2348 egecgiceggaicgecgictecgicggigagicegicggeticgecgagecigagecega 2407
                                                                                                                                                                                                                                                                                                                                                                                            2468 egeegteteegtegeegteaceategeegagteegteteggteteegteaeeateggegt 2527
                                                                                                                                                                                                                                                                                                                                                                                                                                                            2528 cgccgagcccgtcttcgtcaccgagcccgtcttcgtcaccgtcttcgtcgccgagcccaa 2587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ppressor gene; oncogene; antitumour; cytostatic;
CpG dinucleotide; single-nucleotide polymorphism;
                                                                                                                                                                                                                                                                                                                             2408 gcccgagcccgagcagctcgccatcgccgtcgccgtcgccgagctcgagtccatcctcgt
                                                                                                                                                                                                                                                                                                                                                198 CACCATCACCATCACCATCACCATCACCATCACCATCACCATCACCATCACCATCACCAT
                                                                                                                                                                                                                                                                                                                                                                                                               Tumour suppressor gene derived chemically modified sequence #370.
                                                                                                                                                                                             2.7%; Score 76.8; DB 23; Length 651; 56.2%; Pred. No. 1.6e-05;
                                                                                                                                              Sequence 651 BP; 152 A; 86 C; 241 G; 172 T; 0 other;
                                                                                                                 at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                               0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ₹.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Berlin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Human; tumour suppressor gene;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2000DE-1013847.
2000DE-1019058.
2000DE-1019173.
2000DE-1032529.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     15-MAR-2001; 2001WO-EP02955.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2000DE-1043826
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2588 egeegtegtegtegee 2603
                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      cytosine methylation; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ٣
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DNA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ပဲ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (EPIG-) EPIGENOMICS AG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18 CTCCATCTCCATCACC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Piepenbrock
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2001-602752/68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAS46648 standard;
                                                                                                                                                                                                            Local Similarity
Les 144; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          tumour;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200168912-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  06-APR-2000;
07-APR-2000;
30-JUN-2000;
01-SEP-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     15-MAR-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20-SEP-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAS46648;
                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     olek A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          cancer;
                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAS46648/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22222222222
                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        XXX
                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymersae chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in
                                                                                                                                                                                                                                2486 caccatcgccgagtccgtctccgtcaccatcggcgtcgccgagcccgtcttcgt 2545
                                                                                                                                                                                                              2605
                                                                               2426 cgccatcgccgtcgccgtcgccgagctcgagtccatcctcgtcgccgtcgccgt 2485
                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.
                                                                                                                                                                                                          2546 caccgagcccgtcttcgtcaccgtcttcgtcgccgagcccaacgccgtcgtcgtcgcgg
                                                                                                                                                                  New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess
                                                                                                             47143 CGCTGTCGTCGTGCGTGCACTGTGGCCGTGATCGCCACTGTCGTCGCGGCCGGGCCGT
                Length 58857;
                                                 ó
                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DNA encoding novel human diagnostic protein #20760.
                  DB 21;
                                   9.1e-06;
               Score 78.2; DB
Pred. No. 9.1e-
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 1; SEQ ID No 20760; 103pp; English.
                                                                                                                                                                                                                                                                                                                                                                                            BP
                                                                                                                                                                                                                                                                                                                                                                                            AAS84956 standard; cDNA; 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Tang YT;
                2.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       30-MAR-2001; 2001WO-US08631.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2000US-0540217
2000US-0649167
                                                                                                                                                                                                                                                                                                           46963 CGAGCGAGGCGAGGA 46949
                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                           2606 tgtcgggtggggtga 2620
                                Best Local Similarity 62.6
Matches 122; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Drmanac RT, Liu C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2001-639362/73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (HYSE-) HYSEQ INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      P-PSDB; ABG20769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO200175067-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       31-MAR-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        23-AUG-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      biodiversity
                                                                                                                                                                                                                                                                                                                                                                                                                                                           13-FEB-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11-OCT-2001
                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                            AAS84956
                                                                                                                                                                                                                                                                                                                                                                             AAS84956,
                                                                                                                                                                                                                                                                                                         qq
                                                                                 ġ
                                                                                                                                                                                                                                            合
                                                                                                             a
                                                                                                                                              ò
                                                                                                                                                                            셤
                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                            ò
```

SNP;

30-JUN-2000; 2000DE-1032529. 01-SEP-2000; 2000DE-1043826. 02-JUL-2001; 2001WO-EP07537

WO200200928-A2. Homo sapiens.

gene; ds.

03-JAN-2002

```
The invention relates to a nucleic acid comprising a sequence of 18 bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with blaughite, of genes associated with tumour suppression and concogenes having a sequence taken from 516 (actually 533 since numbers 408, 458 and 500 are missing from the sequence listing) sequences. Complementary to (Ss). The nucleic acid may be a peptide nucleic acid coligomer (PNA) of at least 9 nucleotides and may form part of a set of probes for detecting the cytosine methylation state and/or single nucleide polymorphisms and also to be used in an array for analysing diseases associated with CPG dinucleotides e.g. cancers and tumours. The probes can also be used in a method for ascertaining genetic and/or epigenetic parameters for the diagnosis and/or therapy of existing diseases or the predisposition to specific diseases, by analysing cytosine methylations. The parameters may be compared to another set of genetic and/or epigenetic parameters in the diagnosis and/or proposis serving as basis for diagnosis and/or proposis events which are disadvantageous to patients. The present sequence is one of the corresponding odd numbered sequence (e.g. 533 genomic sequences with even numbered Sequence (e.g. complementary sequence of the corresponding odd numbered sequence (e.g. complementary sequence of the corresponding odd numbered sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2362 ccgtccgtcggtgagtccgtcggcttcgccgagcctgagcccgagcccgagcccgagc 2421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2422 agctcgccatcgccgtcgccgaccgagctcgagtccatcctcgtcgccgtctccgtcg 2481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2482 ccgtcaccatcgccgagtccgtctcggtctccgtcaccatcggcgtcgccgagcccgtct 2541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      933 ccGACCCCGACCCCCCCCCCCCCATCCCGATCCTAACTCCCAACCCCGACT 874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         873 CCGACTCCGATCGCGACTCCGATCGCGATCGCGATCGCGATCCTAACCCCGACC 814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      813 CCGACCCCGACTCCGACTCCGACCCCGACTCCAATCGCGACCCCGACCCCGAACC 754
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2542 tegicacegagecegicitegicacegicitegiegecegagecegaegeegiegie 2597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Human; immune system disease; cytosine methylation; antiasthmatic; antiantenic; cytostatic; nootropic; neutianaemic; cytostatic; nootropic; neuroprotective; anti-HTV; anticonvulsant; ophthalmological; antinheumatic; antiarthritic; antidiabetic; antipsoriatic; antinfleumatory; cancer; eye disease; arteriosclerosis; anaemia; acute myeloid leukaemia; Alzheiner's disease; AIDS; epilepsy; neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
analysing diseases associated with cytosine methylation state e.g.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 22; Length 6510;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       the printed specification, but was obtained in electronic mat directly from WIPO at
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     part
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity 56.8%; Score 72.8; DB 22; Length Similarity 56.8%; Pred. No. 8.9e-05; Conservative 0; Mismatches 102; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sèquence 6510 BP; 1437 A; 357 C; 1911 G; 2805 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The sequence data for this patent did not form
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human immune system associated gene SEQ ID NO: 368.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ftp.wipo.int/pub/published_pct_sequences.
                                                                       Claim 1; SEQ ID No 370; 27pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ABL32395 standard; DNA; 6510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ID 2 and ID
is missing)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          134;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ABL32395;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Note:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ABL32395,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid reluxenmia, Alzhamer's disease, Alzhaltepsy, neurofibromatosis, rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel diseases. The present sequence is a gene of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2422 agctcgccatcgccgtcgccgtcgccgagctcgagtccatcctcgtcgccgtctccgtcg 2481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2482 ccgtcaccatcgccgagtccgtctcggtctccgtcaccatcggcgtcgccgagcccgtct 2541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2362 ccgiciccgicggigagiccgicggcitcgccgagccigagcccgagcccgagcccgagc 2421
                                                                                                                                                                                                                                                                                                                                                  Nucleic acid comprising fragment of chemically modified gene, useful for diagnosis and treatment of diseases associated with abnormal cytosine methylation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2542 tegteacegagecegtettegteacegtettegtegeegageeeaaegeegtegte 2597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        933 ccgacccccacccccccccccarccccarccraacccccaacccccaacc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       873 CCGACTCCGATCGCGACTCCGATCGCGATCGCGATCGCGATCCTAACCCCGACC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         813 CCGACCCCGACTCCGACTCCGACCCCGACTCCAATCGCGACCCCGACCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    753 CCGACCCCGACCCCGACCCCGCACCCCGACCCCGAACGCGACGCC 698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 6510;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 6510 BP; 1437 A; 357 C; 1911 G; 2805 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 1; SEQ ID NO 368; 32pp + Sequence Listing; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match 2.5%; Score 72.8; DB 24;
Best Local Similarity 56.8%; Pred. No. 8.9e-05;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0; Mismatches 102;
                                                                                                                                                                                                                                                                                                Berlin K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 134; Conservative
                                                                                                                                                                                                                                                                                              Piepenbrock C,
                                                                                                                                                                                                                                                       (EPIG-) EPIGENOMICS AG
                                                                                                                                                                                                                                                                                                                                      WPI; 2002-130909/17
                                                                                                                                                                                                                                                                                              olek A,
ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

Search completed: July 2, 2002, 13:57:31 Job time: 17153 sec

THIS PAGE BLANK (USPTO)

.

```
4689
13842
38577
88506
444377
444377
28958
28958
28958
28958
28958
28958
28958
28958
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
 US-09-197-649-7/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH: 390
TYPE: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2527
υO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                  2, 2002, 09:11:33 ; Search time 67.49 Seconds
(without alignments)
10441.881 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 7, Appli
Sequence 11, Appl
Sequence 2, Appli
Sequence 173, App
Sequence 1, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 2, Appli
Sequence 1, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1, Appli
12, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 7, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     is the number of results predicted by chance to have a ster than or equal to the score of the result being printed, ived by analysis of the total score distribution.
                                                                                                                                                                                            1 atggatcgttcggagaacat........99cggacacctacctgcagn 2869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Seguence 2, A
Seguence 176,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     182,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sednence Sed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                               /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
/cgn2_6/ptodata/1/ina/5B_COMB.seq:*
/cgn2_6/ptodata/1/ina/6A_COMB.seq:*
/cgn2_6/ptodata/1/ina/6B_COMB.seq:*
/cgn2_6/ptodata/1/ina/PcTUS_COMB.seq:*
/cgn2_6/ptodata/1/ina/PcTUS_COMB.seq:*
              4.5
Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4 US-09103-840A-2

US-09103-840A-2

US-09-103-840A-1

US-09-103-840A-1

US-09-103-649-7

US-09-130-114-2

US-09-130-114-2

US-09-130-114-2

US-09-130-114-2

US-09-130-114-2

US-09-130-114-2

US-09-136-556-173

US-09-136-556-182

US-09-136-556-182

US-09-136-556-182

US-09-025-603-4

US-09-020-603-4

US-09-029-603-4

US-09-029-603-4

US-09-029-603-4

US-09-029-603-4

US-08-596-300A-5

US-08-596-300A-6
                                                                                                                                                                                                                                                                                              of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-312-038-3
US-09-056-556-171
US-09-056-556-167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -08-596-300A-3
                                                                                                                                                                                                                                                                  383533 segs, 122816752 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUMMARIES
              version - 2000
                                                                                                                                                                                                                                                                                                                                                                                                      Listing first 45 summaries
                                                                        sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -09 - 31
                                                                                                                                                                                                                                      Gapop 10.0 , Gapext 1.0
              GenCore
Copyright (c) 1993
                                                                                                                                                                                                                                                                                                                                                                                                                                    Issued_Patents_NA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      score greater than or equal to
and is derived by analysis of
                                                                                                                                                                                                                                                                                                                                                                                        Maximum Match 100%
                                                                        nucleic search, using
                                                                                                                                                                                                                                                                                                                            length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4403765 4
4411529 4
390 4
18596 4
                                                                                                                                                         US-09-917-376-2
2869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     11707 4
                                                                                                                                                                                                                        IDENTITY_NUC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         23673
22293
6416
6416
1294
2064
2064
1081
1081
1181
3718
1129
1289
5389
5389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1931
407
                                                                                                     July
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ........
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Minimum DB seq
Maximum DB seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61.8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61.6
61.6
61.6
61.6
61.6
61
61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100.8
                                                                                                                                                                                                                         Scoring table:
                                                                                                                                                                            Perfect score:
                                                                                                                                                                                                                                                                                                rotal number
                                                                        nucleic
                                                                                                                                                                                             Sequence:
                                                                                                                                                                                                                                                                    Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                    Database
                                                                                                    Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ပ
```

```
sequence 7, Application US/09197649

patent No. 6144550

GENERAL INFORMATION:

APPLICANT: Gold, Larry
APPLICANT: Tuerk, Craig

APPLICANT: Fithow, David

APPLICANT: Smith, Jonathan D.

FILE REFRENCE: NEXOZ/CI-COM

CURRENT APPLICATION NUMBER: US/09/197,649

CURRENT FILING DATE: 1992-01-31

EARLIER PILING DATE: 1992-01-31

EARLIER PILING DATE: 1990-08-02

MUMBER OF SEQ ID NOS: 26

SOFTWARR: Patentin Ver. 2.0

FEARLIER PLOSE

FEARLIER PILING DATE: 1990-08-02

SOFTWARR: PATENTING DATE: 1990-08-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .;
0
                                                                                                                                 19, Appl
1, Appli
6, Appli
1, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gegeegteeggategeegteteegteggtgagteegteggettegeegageetgageeeg 2406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  tegeogagecegtettegteacegagecegtettegteacegtettegtegeegagecea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2407 agcccgagcccgagcagctcgccatcgccgtcgccgtcgccgagctcgagtccatctcg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            tcgccgtctccgtcgccgtcaccatcgccgagtccgtctcggtctccgtcaccatcggcg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sednence sed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OTHER INFORMATION: Description of Artificial Sequence: Sequence; OTHER INFORMATION: having a 120 repeat of ACG flanked by fixed; OTHER INFORMATION: fragments having NCOI restriction sites.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 390;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 100.8; DB 4;
Pred. No. 1.3e-13;
); Mismatches 102;
US-09-105-537-34
US-09-105-537-30
US-09-105-537-30
US-09-105-537-30
US-08-804-227C-7
US-08-804-198-1
US-08-456-837-6
US-08-456-837-6
US-08-457-646A-6
US-08-457-646A-6
US-08-457-646A-6
US-08-457-648A-6
US-08-457-38A-4
US-08-457-38A-4
US-08-457-38A-6
US-08-457-38A-6
US-08-457-38A-6
US-08-457-38A-6
US-08-457-33A-1
US-08-457-33A-1
US-08-457-33A-1
US-08-457-33A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             tch 3.5%;
al Similarity 61.4%;
162; Conservative
```

```
Db 3928586 CCGACCCCGCCTTGCCCCCCGCCGGCGAAGCCGTTGCCGTTTGGGAGGGCG 3928527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3928706 TGGCCGCCGTCACCGCCGTCGGGTCGCCGGGGCCCGGGGCGCCGTCACCACCG 3928647
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3928646 GCCCCGCCGGTGCCGCCGGTGCCGCCGGCCCCGTTGATGCCGCCGGCGCCGGCGTTGCCG 3928587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          positions throughout the sequence g
                                                                                                                                                                                                                                                     2344 ggtgcgccgtccggatcgccgtctccgtcggtgagtccgtcggcttcgccgagcctgagc 2403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2464 tegtegeegteteegtegeegteaceategeegagteegteteggteteegteaceateg 2523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2524 gcgtcgccgagcccgtcttcgtcaccgagcccgtcttcgtcaccgtcttcgtcgccgagc 2583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2584 ccaacgccgtcgtcgtcgccggtgtcgggtggggtgaaggtgcagtataagaataatgat 2643
  3957 CCGACAGTGACGGTTACTGTGACTCCGACACCGACACCAACACGGCGGCGACACCGACA 4016
                                                                                                                                                                                4077 AGCACAAGTTCTATAAGGCCGTGGTTTAAGATAGTGAATGGAGGCAGCAGCAGTGTTGAT 4136
                                                                                                                                                                                                                                                                                                                                                                 4194 GCGGTATGTGACTGGGCACAGATAGGGGCAAGCAATGTGACATTCAATTTTGTGAAGCTG 4253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2404 ccgagcccgagcccgagcagctcgccatcgccgtcgccgtcgccgagctcgagtccatcc 2463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: WHITE, Owen R. APPLICANT: FRASER, Claire M. APPLICANT: PERSER, Claire M. APPLICANT: VENTER, JOHN C. TITLE OF INVENTION: TUBERCULOSIS FOR STRAIN ANALYSIS IN MYCOBACTERIUM TITLE OF INVENTION: TUBERCULOSIS FILE REFERENCE: 24366-20007.00 CURRENT REPERENCE: 24366-20007.00 CURRENT REPLICATION WIDMBER: US/09/103,840A
                                                                                         4017 GGGACACCTGGCACGGGAAGTGGTTTGAAGGTACTATACAAGAACAATGAGACAAGTGCG
                                                                                                                                                                                                                                ttgtcgacggtgacggtgcggtactggttcacccgggatggtggtggctcgtcgacactggtg
                                                 tcgtcgccggtgtcgg---gtggggtgaaggtgcagtataagaataatgattcggcgccg
                                                                                                                                        ggtgataatcagatcaagccgggtttgcaggtggtgaataccgggtcgtcgtcggtggat
                                                                                                                                                                                                                                                                                                                     2773 tacaactgtgactgggcggcgatcgggtgtgggaatatccgcgcctcgttcggctcggtg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 4403765;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 2.8%; Score 79; DB 4; 1
Best Local Similarity 47.2%; Pred. No. 5.8e-08;
Matches 241; Conservative 0; Mismatches 270
                                                                                                                                                                                                                                                                                                                                                                                                              aacccggcgacgccgacggcggacacctacctgcag 2868
                                                                                                                                                                                                                                                                                                                                                                                                                                                           4254 AGCAGCGGAGTGAGTGGAGCGGATTATTACTTGGAG 4289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              at various
a, t, c or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM: Mycobacterium tuberculosis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 2, Application US/09103840A
Patent No. 629428
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, ROBERT D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases
OTHER INFORMATION: represent
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH: 4403765
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-103-840A-2/C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-103-840A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQ ID NO 2
                                                                                                                                                                                                                         2713
                                               2596
                                                                                                                                      2653
                                                                                                                                                                                                                                                                                                                                                                                                                 2833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δλ
                                                                                         a
                                                                                                                                                                              q
                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                           αq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       οg
                                               ö
                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                 Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Treating Cellulose Containing Fabrics Using Truncated Cellulase Enzyme Compositions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2476 ccgtcgccgtcaccatcgccgagtccgtctcggtctccgtcaccatcggcgtcgccgagc 2535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3897 CCGACATCAGCCCCGACACGAGCCCAACAGTGACGCCGACTCCAACGCCGACG 3956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       cogtettogicaccgageccgtettegicaccgtettegicgecgageccaacgccgtcg 2595
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2416 ccgagcagctcgccatcgccgtcgccgtcgccgagctcgagtccatcctcgtcgccgtct 2475
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 11707;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Center, P.O. Box
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              215; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                        INVENTION: Compositions and Methods for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/932,571
FILING DATE: September 19, 1997
ATORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: 1997US001/CIP
TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-540-9200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9.9e-09;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 80; DB 4;
Pred. No. 9.9e-(
0; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: IBM COMPATIBLE
OPERATING SYSTEM: DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,574A
FILING DATE: 19-Aug-1998
CLASSIFICATION: <Unknown>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ..
02
                                                                                                                                                                                                                                                                     APPLICANT: Farrington, Graham K.
Anderson, Paige
Gibbs, Moreland
Bergquist, Peter
Daniels, Roy
Morgan, Hugh W.
                                                                                                                                                                                                   Sequence 1, Application US/09136574A Patent No. 6294366 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH: 11707 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
                                                                                       130 rccrccrccrccrccrccrcc 107
                                                                                                                                                                                                                                                                                                                                                                                                            Williams, Diane P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE DESCRIPTION: SEQ ID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEFAX: 215-540-5818
TELEX: <Unknown>
                                          2587 acgccgtcgtcgtcgccggtgtcg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO: 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCES: 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CITY: Spring House
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              235; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY: USA
ZIP: 19477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ΡA
                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE OF
                                                                                                                                                           SULT 2
-09-136-574A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-136-574A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2536
qq
                                            ò
                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
```

à

```
Db 1188680 ttgccgccggtgccgccgccgccgagcaggccggcgccgccgccgttgccgccgtgaccg 1188739
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Db 1188740 cccgcgtgaccgctgccgccgttaccgccggcggctgaagcccggtcggcgggttggtg 1188799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      OTHER INFORMATION: CDC 1551

OTHER INFORMATION: "n" bases at various positions throughout the sequence
COTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2380 ccgtcggcttcgccgagcctgagcccgagcccgagcccgagcagctcgccatcgccgtcg 2439
                                                                      2527 tegecgagecegtettegteacegagecegtettegteacegtettegtegeegageeea 2586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2440 ccgtcgccgagctcgagtccatcctcgtcgccgtctccgtcgccgtcaccatcgccgagt 2499
225 CCGTTGCCGCCCCCCCCCCCCCCCCCCCCTGA 166
                                                                                                                  165 ccgccgrcgccgccgccgcccrrgccgccgaaaaagccgcrgccgccgcgcgccg 106
                                                                                                                                                            APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TITLE OF INVENTION: TUBBECTUCOSIS
FILE REFERENCE: 24366-20007.00
CORRENT APPLICATION NUMBER: US/09/103,840A
CORRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PATENTIN VET. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2560 tegtcacegtettegtegeegageecaaegeegtegtegtegtegeeggtgtggggtg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2620 aaggtgcagtataagaataatgattcggcgccgggtgataatcagatcaagccgggtttg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 4403765;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 70.2; DB 4; Length 4
Pred. No. 4.9e-06;
0; Mismatches 169; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM: Mycobacterium tuberculosis
                                                                                                                                                                                                                                                                                                                                                                                                                     ; Sequence 2, Application US/09103840A
; Patent No. 6294328
GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 2.4%;
Best Local Similarity 51.4%;
Matches 179; Conservative
                                                                                                                                                                                                                                                           2647 gcgccggg 2654
                                                                                                                                                                                                                                                                                                     45 CCCCCGTG 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQ ID NO 2
LENGTH: 4403765
                                                                                                                                                                                                                                                                                                                                                                         RESULT 5
US-09-103-840A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δŏ
                                                                                                                qq
                                                                                                                                                                                                             g
                                                                      ò
                                                                                                                                                              \dot{Q}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TREAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ó.
                                                                  Db 3928466 cccgcggTTCc6ccgcTTGcCcAcGcCcCCACGCCGCGCGTCGCCGCCGCCGCCG 3928407
                                                                                                                                                                                                                                                  2704 tcggtggatttgtcgacggtgacggtgcggtactggttcacccgggatggtggctcgtcg 2763
                                                                                                                                                                                                          2764 acactggtgtacaactgtgactgggcggcgatcgggtgtgggaatatccgcgcctcgttc 2823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2407 agcccgagcccgagcagctcgccatcgccgtcgccgtcgccgagctcgagtccatcttcg 2466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2467 tegeegteteegtegeegteaceategeegagteegteteggteteegteaceateggeg 2526
                        2644 tcggcgccgggtgataatcagatcaagccgggtttgcaggtggtggataccgggtcgtcg 2703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2347 gegeegteeggategeegteteegteggtgagteegteggettegeegageetgageeeg 2406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            S-09-056-556-176/c
Sequence 176, Application US/09056556
Patent No. 6350456
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skelky, Yasir A.W.
APPLICANT: Dillon, Davin C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND NUMBER OF SEQUENCES: 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          285 AGGCCGCCTGGCCGCCGCCACCGGCAAAGACGCCGTCGCCGCCCGGCCGCCGACACCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 494;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER REDABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFWWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056,556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0; Mismatches 148;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2.5%; Score 71.2; DB 4; 51.9%; Pred. No. 4.2e-07;
                                                                                                                                                                                                                                                                                                CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.457
FELECOMUNICATION INFORMATION:
FELECOMUNICATION INFORMATION:
FELEFAX: (206) 622-6031
INFORMATION FOR SEQ ID NO: 176:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JMBER: US/09/056,556
07-APR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH: 494 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 51.9
Matches 160; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY: USA
ZIP: 98104-7092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-056-556-176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
```

ò g ò pp ò

```
US-09-197-649-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SOFTWARE: P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DNA
  SEQ ID NO 7
                                                                                                                                                                                                                                                                                                                                                                                                          δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: Gold, Larry
APPLICANT: Gold, Larry
APPLICANT: Gold, Larry
APPLICANT: Tuerk, Craig
APPLICANT: Tuerk, Craig
APPLICANT: Pinhow, David
APPLICANT: Smith, Jonathan D.
TITLE OF INVENTION: Systematic Polypeptide Evolution by Reverse Translation
FILE REFERENCE: NEXO2/C1-CON
CURRENT APPLICATION NUMBER: US/09/197,649
CURRENT FILING DATE: 1998-11-23
EARLIER APPLICATION NUMBER: 07/829,461
EARLIER FILING DATE: 1991-08-01
EARLIER FILING DATE: 1991-08-01
EARLIER FILING DATE: 1990-08-02
MUMBER OF SEC ID NOS: 26
SOFTWARE: Patentin Ver. 2.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Db 1188474 ccgccggcgccgtcgagtccggcgatcaggccggcgccaccttgccgccgttgccgccg 1188533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Db 1188714 cccgcgtgaccgctgccgccgttaccgccggcggcttgaagcccggtcggcgggttggtg 1188773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2380 cogtcggcttcgccgagcctgagcccgagcccgagcagctcgccatcgccgtcg 2439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2440 cogtogocgagetcgagtccatcctcgtcgccgtctccgtcgccgtcaccatcgccgagt 2499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2500 ccgtctcggtctccgtcaccatcggcgtcgccgagcccgtcttcgtcaccgagcccgtct 2559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2620 aaggigcagiataagaataatgaticggcgcgggigataatcagaicaagccgggiiig 2679
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                       STRAIN ANALYSIS IN MYCOBACTERIUM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 4411529;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Oy 2680 caggtggtgaataccgggtcgtcgtcggtggatttgtcgacggtgacg 2727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ouery Match 2.4%; Score 70.2; DB 4; Best Local Similarity 51.4%; Pred. No. 4.9e-06; Matches 179; Conservative 0; Mismatches 169;
                                                       GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: FRASER, Claire M.
APPLICANT: FRASER, Claire M.
APPLICANT: VEWTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN
TITLE OF INVENTION: TUBERCULOSIS
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PATENTIN VET. 2.1
SOFTWARE: PATENTIN VET. 2.1
                                                                                                                                                                                                                                                                                                                                                                                ORGANISM: Mycobacterium tuberculosis
COTHER INFORMATION: H37Rv
US-09-103-840A-1
                 Sequence 1, Application US/09103840A Patent No. 6294328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 7, Application US/09197649
Patent No. 6194550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GENERAL INFORMATION:
US-09-103-840A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -09-197-649-7
                                                                                                                                                                                                                                                                                                                                                                  TYPE: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2421 cagetegecategecgtegecgtegecgagetegagtecatectegtegecgtetegte 2480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2481 gccgtcaccatcgccgagtccgtctcggtctccgtcaccatcggcgtcgccgagcccgtc 2540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2661 tcagatcaagccgggtttgcaggtggtgaataccgggtcgtcgtcggtggatttgtcgac 2720
                                                                                                                                                                                                                                                                                                                                       2361 gccgtctccgtcggtgagtccgtcggcttcgccgagcctgagcccgagcccgagcccgag 2420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2541 ttogtcacogagccogtcttcgtcacogtcttcgtcgccgagcccaacgccgtcgtcgtc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PREVENTING, AND TREATING
                                                                                                                                                                                                                                                                                                                                                                                    2601 gccggtgtcgggtggggtgaaggtgcagtataagaataatgattcggcgccgggtgataa
                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                              OTHER INFORMATION: Description of Artificial Sequence: Sequence OTHER INFORMATION: having a 120 repeat of ACG flanked by fixed OTHER INFORMATION: fragments having NCOI restriction sites.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 18596;
                                                                                                                                                                                                                                      Length 390;
                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                         Query Match 2.4%; Score 70; DB 4; Le Best Local Similarity 49.2%; Pred. No. 7.3e-07; Matches 184; Conservative 0; Mismatches 190;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 2.4%; Score 68.8; DB 4; Best Local Similarity 61.9%; Pred. No. 3.1e-06; Matches 109; Conservative 0; Mismatches 67;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 11, Application US/09318448
Patent No. 6210950
GENERAL INFORMATION:
APPLICANT: Johnson, William G.
APPLICANT: Stenroos, Edward S.
TITLE OF INVENTION: METHODS FOR DIAGNOSING, PRI;
TITLE OF INVENTION: DEVELOPMENTAL DISORDERS
FILE REFERENCE: 601-1-057
CURRENT APPLICATION NUMBER: US/09/318,448
CURRENT FILING DATE: 1999-05-25
NUMBER OF SEQ ID NOS: 46
LENGTH: 390
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            : PatentIn Ver. 2.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2721 ggtgacggtgcggt 2734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        363 gacgacgccatggt 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; ORGANISM: Homo sapiens
US-09-318-448-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-318-448-11/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 18596
```

අ

ŏ

```
2348 egeegteeggategeegteteegteggtgagteegteggettegeegageetgageeega 2407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2408 gcccgagcccgagcagctcgccatcgccgtcgccgtcgccgagctcgagtccatcctcgt 2467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2468 egeogicicogicgcegicaccaicgcegagicegicteggicicegicaceaicggegi 2527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2528 egecgagecegtettegteacegagecegtettegteacegtettegtegeegageceaa 2587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        183 cégcécccartaacaégccéccacéccéccégccécceccercéccéccéccéccca 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             243 CGCGCCACCGTCGACACCGGCACCGCCGCCGCCAGTGCCACCGTGCCCGCCGGCGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 67.8; DB 4; Length 4
Pred. No. 2.2e-06;
0; Mismatches 112; Indels
COMPUTER: Floppy disk COMPUTER: TBM PC COMPATIBLE OPPRATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: PATENTIN Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056,556 CLASSIPTING OF APPLICATION NUMBER: US/09/056,556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 1, Application US/09773816

Patent No. 6340774

GENERAL INFORMATION:
APPLICANT: Stanford University
APPLICANT: Stanford University
TITLE OF INVENTION: NON-STEROIDAL ESTROGEN-RECEPTOR
TITLE OF INVENTION: NON-STEROIDAL ESTROGEN-RECEPTOR
TITLE OF INVENTION: ANTAGONISTS
FILE REPERENCE: 28600-20210.00
CURRENT APPLICATION NUMBER: US 60/243,458
PRIOR FILING DATE: 2000-10-25

PRIOR FILING DATE: 2000-01-31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NUMBER OF SEQ ID NOS: 1
SOFTWARE: FastSEQ for Windows Version 4.0
                                                                                                                                                                                                                                                   ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.457
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
                                                                                                                                                                                                                                                                                                                                                                                           TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 1
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH: 407 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 2.48
Best Local Similarity 54.78
Matches 135; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2588 cgccgtc 2594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         123 AGCCGCC 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: DNA
ORGANISM: Human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ;
US-09-056-556-173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     us-09-773-816-1/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQ ID NO 1
LENGTH: 23673
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TREAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                     2348 egecgiceggaiegecgietecgieggigagieegieggetiegeegaageeigageeega 2407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2408 gcccgagcccgagcagctcgccatcgccgtcgccgtcgccgagctcgagtccatcctcgt 2467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2528 cgccgagcccgtcttcgtcaccgagcccgtcttcgtcaccgtcttcgtcgccgagcccaa 2587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2468 egcegtetecgtegeegteaceategeegagteegteteggteteegteaceateggegt 2527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0; Gaps
                                                                                                                                                              10271 CCATCACCATCACCATCACCATCACCATCACCATCACCATCACCATCACCATCACC
                                                                                                                                   2548 ccgagcccgtcttcgtcaccgtcttcgtcgccgagcccaacgccgtcgtcgcc 2603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 173, Application US/09056556

Batent No. 6350456

GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
ATTLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND NUMBER OF SEQUENCES: 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 68.4; DB 2; Length 1931;
Pred. No. 2.3e-06;
0; Mismatches 121; Indels 0
                                                                                                                                                                                                                                                                                                                                                                APPLICANT: Horilck, Robert A. APPLICANT: Damaj, Bassam B. APPLICANT: Damaj, Bassam B. APPLICANT: Robbins, Alan K. TITLE OF INVENTION: Bukaryotic Cells Stably Expressing TITLE OF INVENTION: Eukaryotic Cells Stably Expressing TITLE OF INVENTION: From Multiple Transfected Episomes FILE REFERENCE: 0867/1D903US1
CURRENT APPLICATION NUMBER: US/09/130,114
CURRENT FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
                                                                                                                                                                                                                                                                                               Sequence 2, Application US/09130114 Patent No. 5976807 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2588 cgccgtcgtcgtcgccggtgtc 2609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 2.45
Best Local Similarity 53.85
Matches 141; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CITY: Seattle
STATE: Washingt
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98104-7092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-056-556-173/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; TYPE: DNA
; ORGANISM: EBNA
US-09-130-114-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENGTH: 1931
```

g

õ g ò

g ò g

ò g

nucleic acid EDNESS: double

linear

```
STRANDEDNESS:
                                         TOPOLOGY: 11 MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                    фQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                    ;
0
                                                                                                                                                                                                            2342 giggigacogicogalogacogicogicogicogicogicogicogicogocogagociga 2401
                                                                                                                                                                                                                                        7729 GGGGCTGGCGACAAACTGGGTCTCCCGGTCACGTTGGTCGTCGACACCCCGGGGGCGTTA 7670
                                                                                                                                                                                                                                                                                           2402 gecegagecegagecegageategecategeegtegeegtegeegagetegagteeat 2461
                                                                                                                                                                                                                                                                                                                                                                            2462 cctcgtcgccgtccgtcgccgtcaccatcgccgagtccgtctcggtctccgtcaccat 2521
                                                                                                                                                                                                                                                                                                                                                                                                         eggegtegeegageeegtettegteacegageeegtettegteacegtettegtegeega 2581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       7549 GGGGCTCGCGCTCGCCGTCGCCAACCGTGTCCTGATGTTCGAGAACGGCATCTACTCGGT 7490
                                                                                                                                                                                                                                                                                                                                     7669 CCCGGCGCGGGCCGAGGAGCAGGCCAGGCCGTCGCCATCGCGGAGAACCTGCGCCT
                                                                                                                                                                    .;
0
                                                                                                                      Length 23673,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE OF INVENTION: METHOD FOR INCREASING TITLE OF INVENTION: THERMOSTABILITY IN CELLULASE ENZYMES NUMBER OF SEQUENCES: 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2582 goccaacgccgtcgtcgtcgccggtgtcgggtggggtgaaggtgcag 2628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7489 CATCAGCCCCGAGGGCTGCGCGGCCATCCTGTGGCAGGACCCCGCCG 7443
                                                                                                                      Score 67.8; DB 4; Length 2
Pred. No. 5.4e-06;
0; Mismatches 137; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ADDRESSEE: National Renewable Energy Laboratory
STREET: 1617 Cole Boulevard
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICATION NUMBER: US/08/604,913B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 12, Application US/08604913B Patent No. 5712142
GENERAL INFORMATION:
APPLICANT: Adney, William S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER: IBC PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: ASC II (DOS) text CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/276,213
FILLING DATE: 15-070-1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         95-26
               ; LOCATION: (1)...(23623)
; OTHER INFORMATION: n = A,T,C or G
US-09-773-816-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME: Edna M. O'Connor
REGIESTRATION NUMBER: 29,252
REFERENCE/DOCKET NUMBER: 95
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: Adney, William S. APPLICANT: Thomas, Steven R. APPLICANT: Himmel, Michael E. APPLICANT: Baker, John O. APPLICANT: Chou, Yat-Chen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        INFORMATION FOR SEQ ID NO: 12: SEQUENCE CHARACTERISTICS:
                                                                                                                        2.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              303/384-7573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2293 base pairs
                                                                                                                                            Best Local Similarity 52.3
Matches 150; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  303/384-7499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CORRESPONDENCE ADDRESS:
NAME/KEY: misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY: U.S.A. ZIP: 80401-3393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDALL
STREET: 101.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 12
US-08-604-913B-12
                                                                                                                        Query Match
                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                  Dp
                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

```
Bergquist, Peter
Daniels, Roy
Morgan, Hugh W.
Williams, Diane P.
TITLE OF INVENTION: Compositions and Methods for
Treating Cellulose Containing Fabrics Using Truncated
Cellulase Enzyme Compositions
                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                             2086 GTCGGCGAGTCGGACGCCGACGCCTACTCCGACGCCGACGCCGACGCCAACGCT 2145
                                                                                                                                                                                                                                                                                                                                                                                                  2550 gagcccgtcttcgtcaccgtcttcgtcgccgagcccaacgccgtcgtcgtcgccggtgtc 2609
                                                                                                                                                                                                                                                                                                                                                                                                                                           2146 GACCCCTACTGCTACGCCCACGCCAAGCCCGACGCCGTCACCGACGCCACCCTC 2205
                                                                                                                                                                                                                                                  2026 ATCTGCATCGCCTAGCAGTCAACCGTCCGGTGTCGCCGTCTCCGTCGCCGAGCCC 2085
                                                                                                                                                                                                    2430 atcgccgtcgccgtcgccgagctcgagtccatcctcgtcgccgtctccgtcgccgtcacc
                                                                                                                                                                                                                                                                                                   2490 atcgccgagtccgtctccgtctccgtcaccatcggcgtcgccgagcccgtcttcgtcacc
                                                                                                                                                      ö
                                                                                                      Length 2293;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE: Howson and Howson STREET: Spring House Corporate Center, P.O. Box 457 CITY: Spring House
                                                                                                                                                      Indels
                                                                                                 Score 66.6; DB 1;
Pred. No. 5.9e-06;
0; Mismatches 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE/DOCKET NUMBER: 1997US001/CIP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,574A
FILIM DATE: 19-Aug-1998
CLASSIFICATION: -(UKNOWN>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/932,571
FILING DATE: September 19, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: Farrington, Graham
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 2, Application US/09136574A Patent No. 6294366 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER: IBM Compatible OPERATING SYSTEM: DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Anderson, Paige
Gibbs, Moreland
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEFAX: 215-540-5818
                                                                                                   Query Match 2.3%;
Best Local Similarity 60.0%;
Matches 111; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO: 2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NUMBER OF SEQUENCES: 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY: USA
       E1-CAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ZIP: 19477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STATE: PA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2610 gggtg 2614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2206 CGGAG 2210
; NAME/KEY:
US-08-604-913B-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-136-574A-2
```

.; 0

```
TELEPHONE:
                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                       5
                                                                                                                                                                                                                                                                                                                                                          2552 gecegicticgicacegicticgicgeegageecaaegeegicgicgicgeeggigicgg 2611
                                                                                                                                                                                                                                                                                                                          2600 CAGTGACAGCAACGCCGACGCCGACTCCTACACCGACACGGGGGTCACCTGGTACGG 2659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   cggcgatcgggtgtgggaatatccgcgcctcgttcggctcggtgaacccggcgacgcga 2848
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2492 cgccgagtccgtctcggtctccgtcaccatcggcgtcgccgagcccgtcttcgtcaccga 2551
                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                           2669 agccgggtttgcaggtggtgaataccgggtcgtcgtcgtggtggatttgtcgacggtgacgg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                tgcggtactggttcacccgggatggtggctcgtcgacactggtgtacaactgtgactggg
                                                                                                                                                                                       9
                                                                                                                                                      Length 6416;
                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: Broadway, Roxanne M.
APPLICANT: Braman, Gary E.
TITLE OF INVENTION: FUNGUS AND INSECT CONTROL WITH
TITLE OF INVENTION: CHITINOLYTIC ENZYMES
CORRESPONDENCES: 4
CORRESPONDENCE ADDRESS:
                                                                                                                                                    2.3%; Score 66.4; DB 4;
52.1%; Pred. No. 8.2e-06;
tive 0; Mismatches 176;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP STREET: Clinton Square, P.O. Box 1051
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATTORNEY AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/20120
TELECOMMUNICATION INFORMATION:
                                 TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/025,691
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           E: Floppy disk
IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
SEQUENCE CHARACTERISTICS:
LENGTH: 6416 base pairs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 2, Application US/09025691
Patent No. 6069299
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2849 cggcggacacctacctgcag 2868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2897 GAGCGGATTATTACCTGGAG 2916
                                                                                                                                                    Query Match 2.39
Best Local Similarity 52.19
Matches 198; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER: IBM PC OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New York
: U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STREET: Clinton
CITY: Rochester
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY: U
ZIP: 14603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SOFTWARE:
                                                                                                    US-09-136-574A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-025-691-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
```

```
2348 egecgiceggaicgecgicteegicggigagicegicggeticgecgagecigagecega 2407
                                                                                                                                                                                                                                                                                                                                                                                                                                                  2408 gcccgagcccgagcagctcgccatcgccgtcgccgtcgccgagctcgagtccatcctcgt 2467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2468 egeegteteegtegeegteaceategeegagteegteteggteteegteaceateggegt 2527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       589
                                                                                                                                                                                                                                                                                                                                                                                     350 GGCGGCCCAGGCCGCCCGGCCGTACCGGCCGCGGGGACCGGTTACTGGCAGAA 409
                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   410 crrcaacaacececeacerecaeacerecereeceaeereceeaeeeraeearear
                                                                                                                                                                                                                                                                    Length 1294;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-276-213-6
Sequence 6, Application US/08276213
Sequence 6, Application US/08276213
Patent No. 553655
GENERAL INFORMATION:
APPLICANT: Laymon, Robert
APPLICANT: Himmel, Michael
TITLE OF INVENTION: GENE ENCODING FOR THE E1 ENDOGLUCANASE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
CORRESPONDENCE ADDRESS:
STREET: 1617 Cole Boulevard
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                    2.3%; Score 66.2; DB 3;
54.3%; Pred. No. 6.4e-06;
                                                                                                                                                                                                                                                                                                                0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME: O'Connor, Edna
REGISTRATION NUMBER: 29,252
REFERENCE/POCKET NUMBER: NREL IR# 94-08
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303)231-1000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER: US/08/276,213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ZIP: 80401-3393

COMPUTER READABLE FORM:
MEDUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-
(716) 263-1304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEPHONE: (303)231-1000
TELEFAX: (303)231-1098
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                      54.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ATTORNEY/AGENT INFORMATION:
                                                              SEQUENCE CHARACTERISTICS:
LENGTH: 1294 base pairs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3004 base pairs
                    TELEFAX: (716) 263-1600 INFORMATION FOR SEQ ID NO: 3
                                                                                                                                                                                                                                                                                           Best Local Similarity 54.3
Matches 134; Conservative
                                                                                                          TYPE: nucleic acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CLASSIFICATION: 435
                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2588 cgccgtc 2594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             590 CGACGCC 596
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Golden
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     USA
                                                                                                                                                    ; TOPOLOGY: Ii
; MOLECULE TYPE:
US-09-025-691-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ပ္ပ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STATE:
                                                                                                                                                                                                                                                                      Query Match
```

```
;
0
                                                                                                                                                               Gaps
                                                                                                                    Query Match 2.3%; Score 66; DB 1; Length 3004; Best Local Similarity 59.7%; Pred. No. 8.5e-06; Matches 111; Conservative 0; Mismatches 75; Indels
                                                                                                                                                                                                                                                                                                                                                                             Search completed: July 2, 2002, 16:53:17 Job time: 27704 sec
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
MANIT-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-276-213-6
                                                                                                                                                                                                                                                                                                        2609 cgggtg 2614
                                                                                                                                                                                                                                                                                                                               2205 CCGGAG 2210
                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                 QΩ
                                                                                                                                                                                      g
                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                  Q
```

71.6 2.5 522 9 AIG01018 71.6 2.5 606 9 AIG01018 71.2 2.5 506 9 AIG01018 71.2 2.5 587 10 C795374 71 2.5 259 12 AZ374273 70.2 2.4 809 12 CNS02CFH 70 2.4 490 10 BE494661 69.8 2.4 490 10 BE627594 69.2 2.4 457 00 BE627594 69.2 2.4 696 9 AL506972 69.2 2.4 696 9 AL506972 69.2 2.4 696 9 AL506972 68.6 2.4 365 12 L2644X	C 31 68.4 2.4 557 10 BF14396 BF145396 WHEIB41-1 C 32 68.2 2.4 557 10 BF145396 WHEIB41-1 C 34 68.2 2.4 62.9 9 AU496745 AV466745 EXPRM041 C 34 68.2 2.4 62.9 10 BF429394 BF429394 AU197994 AU19799 AU197994 AU	DNA linear GSS 30-NOV Rattus norvegicus genomic c ata; Vertebrata; Euteleosto ognathi; Muridae; Murinae;	D ECORI segment O, USA CHORI-230 P BAC library (pdejong@mail.cho.orges mation.htm). BAC end oac_end_intro.html
GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd. OM nucleic - nucleic search, using sw model Run on: July 2, 2002, 09:11:33; Search time 2155.46 Seconds (without alignments) 17itle: US-09-917-376-2 Perfect score: 2869	table:	Database: EST:* 1: em_estba:* 2: em_esthum:* 3: em_esthum:* 5: em_estruc:* 5: em_estroc:* 6: em_estroc:* 7: em_estroc:* 9: gb_est:* 10: gb_est2:* 11: gb_htc:* 12: gb_gss:* 13: em_gss_inv:* 14: em_gss_inv:* 15: em_gss_inv:* 16: em_gss_inv:* 17: em_gss_inv:* 18: em_gss_inv:* 18: em_gss_inv:* 19: em_gss_inv:* 19: em_gss_inv:* 10:	S derived by analysis of the total score distribution. Summaries Description BH306809 Substance Substance Substance Substance Substance Summaries Summaries Summaries Summaries Summaries Summaries Summaries Summaries Substance Substance

```
AZ928926 464 bp DNA linear GSS 01-APR-2001 479.dlf19912.sl Saccharomyces kluyveri Saccharomyces kluyveri genomic clone 479.dlf19912.sl, DNA sequence.
                                                                                                                                                                                                                                                       5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota; Fungl: Asconycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetales; Saccharomycetales; Lough and Johnston, L. (Bases I to 464).

Cliften, P. F., Hillier, L.W., Fulton, L., Graves, T., Miner, T., Gish, W.R., Waterston, R.H. and Johnston, M. Surveying Saccharomyces genomes to identify functional elements by comparative DNA sequence analysis
Unpublished (2001)
Contact: Johnston M
                                                                   /cell_type="Brain"
/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
CHORI-230 Rat (BN/SSNHSd/MCW) BAC library produced by
Pieter de Jong"
                                                                                                                                                                                                                                                                                                                                                               2468
                                                                                                                                                                                                                                                                                                                                                                                                                                 2350 ccgtccggatcgccgtctccgtcggtgagtccgtcggcttcgccgagcctgagcccgagc 2409
                                                                                                                                                                                                                                                                                                                                                            268 CGGTCTCTGTCTCTGTCTCTGTCTCTGTCTCCGTCTCCGTCTCCGTCTTCCGTCTT 327
                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                     2;
                                                                                                                                                                                                               Score 93.4; DB 12; Length 559;
Pred. No. 7.6e-05;
0; Mismatches 86; Indels 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Department of Genetics
Washington University Medical School
Washington University Medical School
Box 822, 4566 Scott Ave., St. Louis, MO 63110, USA
Tel: 314 362 2735
Fax: 314 362 7855
Email: mj@genetics.wstl.edu
Class: random plasmid subclone.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /organism="Saccharomyces kluyveri"
/strain="NRRL Y-12651 (CBS 3082)"
/db_xref="taxon:4934"
/clone="479.dif19912.s1"
/clone=11b="Saccharomyces kluyveri"
/note="Random genomic sequence"
a 69 c 163 g 113 t
           /clone="CH230-100P7"
/clone_lib="CHORI-230 Segment 1"
/sex="Female"
                                                                                                                                            190 t
/db_xref-"taxon:10116"
                                                                                                                                          65 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AZ928926.1 GI:13499832
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Saccharomyces kluyveri.
                                                                                                                                                                                                                 tch 3.3%;
al Similarity 65.5%;
167; Conservative
                                                                                                                                          232 C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2589 gecgtegtegtegee 2603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    507 TCCGTCTCCGTCTCC 521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             T. .464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ø
                                                                                                                                                                                                                 Query Match
Best Local S
Matches 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BASE COUNT
ORIGIN
                                                                                                                                            BASE COUNT
DRIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AZ928926/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SOURCE
                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                        Db
                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

```
/organism="Neurospora crassa"
/strain="74-0R23-IV A (FGSC 2489)"
/db_raref="taxon:5141"
/clone="swally"
/clone="swally type A"
/tissue_type="Mycelium"
/dev_stage="Mycelium"
/de
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BF072664 ST 18-OCT-2000 NCSM3H7T3 Subtracted Mycelial Neurospora crassa cDNA clone SM3H7 5' similar to avicelase III, Aspergillus aculeatus, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Neurospora crassa.

Neurospora crassa.

Neurospora crassa.

Surdariales, Ascomycota; Pezizomycotina; Sordariomycetes;

Sordariales; Sordariaceae; Neurospora.

1 (bases 1 to 546)

Nelson M.A., Kang.S., Braun,E.L., Crawford,M.E., Dolan,P.L.,

Leonard,P.M., Mitchell,J., Armijo,A.M., Bean,L., Blueyes,E.,

Cushing,T., Errett,A., Fleharty,M., Gorman,M., Judson,K., Miller,R.,

Ortega,J., Pavlova,I., Perea,J., Todisco,S., Trujillo,R.,

Valentine,J., Wells,A., Werner-Washburne,M., Yazzie,S. and Natvig
                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2485 tcaccatcgccgagtccgtctcggtctccgtcaccatcggcgtcgccgagcccgtcttcg
                                                                                                                                               2425 tegecategeegtegeegtegeegagetegagtecatectegtegeegtegeeg
                                                                     ó
       Length 464;
                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Contact: Natvig, D.O./Nelson, M.A.
Department of Biology
University of New Mexico
Castetter Hall, Albuquerque, NM 87131, USA
Tel: 505 277 3411
Fax: 505 277 0304
Score 92.2; DB 12;
Pred. No. 0.00011;
0; Mismatches 58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Neurospora crassa
Fungal Genet. Biol. 21, 348-363 (1997)
97435549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ngp@biology.unm.edu.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BF072664.1 GI:10866169
3.2%;
ilarity 68.6%;
Conservative C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1. .546
                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2605 gtgtc 2609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     143 TCGTC 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              97
                                        Best Local Sim
Matches 127;
       Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BASE CÓUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JOURNAL
MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
BF072664
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMMENT
                                                                                                                                                                                                                                                                                           δλ
                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Dp
```

m

```
Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Lansing, Mi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sedneuce.
                                                                                            65
                                                                                                                                                                     Query Match
                                                                                                                                                                                        Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEFINITION
                                                                                            BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AA720369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOURCE
                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                 à
                                       5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CNSV/HUW 306 bp DNA linear GSS 02-OCT-2001 Anopheles gambiae GSS SP6 end of clone 24A06 of library Notrebamel from strain PEST of Anopheles gambiae (African malaria mosquito), 91610ARE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Direct Submission

Submitted (01-0CT-2001) Genoscope - Centre National de Sequencage :

BP 191006 EVRX cedex - FRANCE (E-mail : seqrefégenoscope.cns.fr - Web : www.genoscope.cns.fr)

( bases 1 to 306)

Roth,C.W., Brey,P.T., Ke,Z. and Collins,F.H.

Direct Submission

Submitted (01-0CT-2001) BBMI, Institut Pasteur, 25, rue du Dr.
                                                                                                                                               1326 ctacggaacaggcgcgacgttgtacgcaacaatgatctcacgaagtgggactccggcgg 1385
                                                                                                                                                                                                                        1386 ccagattcatatcgcgccgatggtcaaaggattggaggaggaggcggtaaacgatctcat 1445
                                                                                                                                                                                                                                                                                                                                                                           ----cccacgccgacgttactgccgtgccatcgacgatttcacgtcaccggtgttcac 1556
                                                                                                                                                                                                                                                                                                                                                                                                                                                      gaccggcaccagcgtcgactatgcggaattgaatccgtcgatcatcgttcgcgctggaag 1616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1617 tttcgatccatcgagccaaccgaacgacaggcacgtcgcgttctcgacagacggcggcaa 1676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1677 gaactggttccaaggcagcgcacctggcggggtgacgacgggcggcaccgtcgccgcatc 1736
                                                                        1266 gaageteggeteggatggatgaagegatggeaategateegtteaactetgateggatget 1325
                                                                                                                                                                                                                                                                                                 1446 cagcccgccgtctggcgccccgctcatcagcgctctcggagacctcggcggcttca---- 1501
                                                                                                                                                                                                                                                                                                                                      291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        352 CAGCTCCGCGGACGTCGACTACGCCGGCAACGTGCTGCGCATCGCGCATCGGC--- 408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             409 -------CTCCGGCTCCGGCGCGCAGCAGCGGCCGTATCCTCGGACGCGGCGCC 456
                                                                                                52 GCACCTCGGCTGGATGATCGAATCCCTCGAGATCCCCCTCGACACCGCCACCATGGCT 111
                                                                                                                                                                                  112 CTACGCCACCGGTCTCACTGTTTACGGCGGCCACGACCTGACCAAGTGGGACACCGTCCA 171
                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
Culicoidea; Anopheles.
1 (bases 1 to 306)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Roux, Paris 75015, France
This close is from an A. gambiae BAC library provided by F.H.
Collins and sequenced by Genoscope in collaboration with the
Laboratory of Biochem. and Biol. Molec. of Insects, Institut
                                                                                                                                                                                                                                                   172 CAACGIGACCAICCAGCCTIGGCCGTIGGCATCGAAGAAATGGCTGTCCTCGGTCTCGC
                                                                                                                                                                                                                                                                                                                        292 CAAGTCCAGCTCCGACCTCGGCACCTCTCCCAAGACACCCTGGATGACGCCCCAGTGGGC
                                       24;
 Length 546;
Score 90; DB 10; Length 54
Pred. No. 0.00024;
0; Mismatches 220; Indels

    .306
    /organism="Anopheles gambiae"

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1737 ggccgacggctctcgtttcgtctgggctcc 1766
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  517 TGCCGACGCCGACACCATCGTCTGGTCTTC 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    African malaria mosquito.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AL610466.1 GI:15916651
3.18;
                                       Conservative
               Local Similarity
es 266; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Genoscope
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AL610466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pasteur.
 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM
                                     Matches
                                                                                                                                                                                                                                                                                                                                                                           1502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                  1557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CNS07H0W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
                                                                                                                                                 ò
                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

```
/clone_lib_minanda-PRL2"
/clone_lib_minanda-PRL2"
/clone_lib_minanda-PRL2"
/clone_lib_minanda_PRL2"
/clone=Vector: lambda zip-Lox; Site_l: Sal; Site_2: Not;
Lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark-rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
sliques. The vector is BRL's lambda Zip-Lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dT primed cDNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   thale cress.

Arabidopsis thaliana
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 236)
Newman,T., GbEruijn,F.J., Green,P., Reegstra,K., Kende,H., McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel,E. and Somerville,C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones Pblant Physiol. 106, 1241-1255 (1994) 95148729
                                                                                                                                                                                                                                                                                                                                       ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EST 30-DEC-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AA720369 236 bp mRNA linear EST 30-DEC-1'
33562 Lambda-PRL2 Arabidopsis thaliana cDNA clone 171N14T7, mRNA
                                                                                                                                                                                                                                                                                                                                                                                                    2425 tegecategecegtegecegtegecegagetegagtecatectegtegecegtetecegtegecg 2484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 180 TCACCGTCACCGTCACCGTCACCGTCACCGTCACTGTTACCGTCACCGTCACCG 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                              120 TCACCGTCACCGTCACCGTCACCGTCACCGTCACCGTCACCGTCACCGTCACCGTCACCG 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Contact: Thomas Newman
MSU-DDE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg.,E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2485 teaceategeegagteegteteggteteegteaceateggegtegeegageeegtetteg
                                                                                                                                                                                                                                                                                                                                       ó;
                                                                                                                                                                                                                                                                             Length 306;
                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 84.8; DB 12;
Pred. No. 0.0014;
0; Mismatches 47;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /organism="Arabidopsis thaliana"
/Astrain="var columbia"
/Ab_xref="taxon:3702"
/clone="171N14T7"
                                                                                                                                                      ų
                                                                                                                                                      89
/strain="PEST"
/db_xref="taxon:7165"
/clone="24006"
/clone=lib="NotreDame1"
/note="end : SP6"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Email: 22313tcn@ibm.cl.msu.edu
Seq primer: T7 dye primer.
Location/Qualifiers
                                                                                                                                                             ь
                                                                                                                                                      63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AA720369.1 GI:2733979
                                                                                                                                                                                                                                                                             3.0%;
                                                                                                                                                      110 c
                                                                                                                                                                                                                                                                                                                                       Matches 113; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Tel: 517-353-0854
Fax: 517-353-9168
```

```
poly(A)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Secale cereale
                                                                                        Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BF145280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                        á
                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /organism="Secale cereale"
/cultivar="Blanco"
/cultivar="Blanco"
/db.refe="texnon:4550"
/clone="While156_1607_N13"
/clone="While156_1607_N13"
/clone="While156_1607_N13"
/clone="While156_1607_N13"
/dev_stage="Anther"
/dev_stage
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       478 bp mRNA linear EST 02-AUG-2000 WHE1261_G07_N132S Secale cereale anther cDNA library Secale cereale cDNA clone WHE1261_G07_N13, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Sperimatophyta; Magnollophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Secale.

1 (bases 1 to 478)
Anderson,0.0.; Butler,B., Chao,S., Choi,D.W., Close,T.J., Fenton A.B., Gustafson,J.P., Han,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Ross,K., Seaton,C.L. and Tong,J.C. The structure and function of the expressed portion of the wheat Unpublished (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
                                                                                                                                                                             ö
                                                                                                                                                                                                                                   2457 tecatectegtegeegteteegtegeegteaceategeegagteegteteggteteegte 2516
                                                                                                                                                                                                                                                                                                                                                2517 accatoggogtogcogagocogtottogtoacogagocogtottogtoacogtottogto 2576
                                                                                                                                                                                                                                                                                                                                                                                        61 TCCGTCTCCGTCTCCGTCTCCGTCTCCGTCTCCGTCTCCGTCTNCGTCTCCGTC 120
                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                  Fax: 5105595818
Email: oandersn@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
sequence have been trimmed score less than 20
seq primer: Stratagene SK primer
Location/Qualifiers
                                                                                                                                                                             ö
                                                                                                                   Length 236;
   others
                                                                                                                                                                          40; Indels
σ
                                                                                                             Score 81.6; DB 9;
Pred. No. 0.0043;
   ı,
                                                                                                                                                                          0; Mismatches
71
   6
   46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2577 gccgagcccaacgccgtcgtcg 2598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121 TCCGNATCCATNTCTNTNTCG 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BE496065
BE496065.1 GI:9662658
                                                                                                                2.8%;
71.8%;
   O
                                                                                                                                                                          102; Conservative
79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Secale cereale
                                                                                                                                              Similarity
   α
31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     51
                                                                                                                   Query Match
                                                                                                                                                 Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VERSION
KEYWORDS
SOURCE
ORGANISM
BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEFINITION
                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
BE496065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMMENT
                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                       유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
/organism="Secale cereale"
/cultivar="Blanco"
/db_xref="taxon:4550"
/db_cref="taxon:4550"
/clone="WHE1834_fi0_k20"
/clone=Lype="Anther"
/tissue_type="Anther"
/dev_stage="Adult plant before anthesis"
/lab_host="E. col! SoLR"
/lab_host="E. col! SoLR"
/lab_host="E. col! SoLR"
/lab_host="Tamoda Uni-ZAP XR, excised phagemid:
/lab_host="Tamoda Uni-ZAP XR, excised phagemid:
site_l: EcoRI: Site_l: XhOI: Plants were grown in the
greenhouse. Anthers were harvested and pooled from early
meiosis to late melosis. The tissue, total RNA, and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WHE1834_f10_k20zS Secale cereale anther cDNA library Secale cereale con constitution with the constitution with the constitution of the constitution with the constitution with the constitution of the constitution with the cereal cereal constitution of the constitution of the constitution of the constitution of the cereal constitution of the con
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnollophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Secale.

7 Triticeae; Secale.

1 (bases 1 to 523)
Anderson,O.D., Butler,E., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Gustafson,J.P., Han,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Ross,K., Seaton,C.L. and Tong,J.C.

7 The structure and function of the expressed portion of the wheat genomes - Anther CDNA library from rye
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RNA were prepared (Butler, Ross and Gustafson) at
                                                                                                                                                                                                                                                                                                                                                                2363 cgtctccgtcggtgagtccgtcggcttcgccgagcctgagcccgagcccgagcccgagca 2422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2483 egicaccaiogocgagicogicioggiciocgicaccaioggogicgocgagocogicii 2542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2543 cgtcaccgagcccgtcttcgtcaccgtcttcgtcgccgagcccaacgccgtcgtcgtcgc 2602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     225 CATCCGCGCGCGCACCTCGCTCGCCATCGCCGCCTCGGCCCGCTCAGTGGCC 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            345 cerecedecerecedadadadecedececedadecederecarracedecereacerer 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      405 CGTCCCCGACCCCCCCCCCCCCCCCCCCCCCCACAGGTTCTCCGCCCCACAGGTTCTCCGCCCCACAGGTTCTCCGCCCCACAGGTTCTCCGCCCACAGGTTCTCCGCCCCACAGGTTCTCCGCCCCACAGGTTCTCCGCCCACAGGTTCTCCGCCCACAGGTTCTCCGCCCACAGGTTCTCCGCCCACAGGTTCTCCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTCACAGGTTC
                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Email: oandersn@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                    Indels
         10;
                                                                                                                                                                                    0; Mismatches 108;
Score 81.2; DB Pred. No. 0.005;
         2.8%;
                                                                                                                                                                                    Matches 146; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    cggtgtcgggtggg 2616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      465 CGTGGCCGGCAGCG 478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Fax: 5105595818
```

```
"Iodx
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               house mouse.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       308, b.
USA
                            408 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM
                            BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AZ386138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE
                                             ORIGIN
                                                                                                                                                                                                                                                                             δλ
                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 (bases 1 to 1315)
SongH., PengY., Gao,G., Xiao,H., Xu,X., Li,N.,
SongH., PengY., Gu,Y., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N.,
Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu
,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Ye,M., Zhang,Q.,
Han,Z., Chen,Z., Hu,R. and Chen,J.
Homo sapiens NPD library cDNA clones
Unpublished (2000)
Contact Qinghua Zhang
Shanghai Institute of Endocrinology, Rui-Jin Hospital
197 Rui-Jin II Road, Shanghai 200025, P. R. China
Tel: 86-21-64743206.
Email: mbshi@ms.stn.sh.cn
University of Missouri, Columbia. A cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript phagemids in the TJ Close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors). "
                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EST 19-OCT-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1315)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /tissue_type="pituitary"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AV752548 NPD Homo sapiens cDNA clone NPDAZH01 5', mRNA sequence. AV752548 NPD (I:10910396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                2483 cgtcaccatcgccgagtccgtctcggtctccgtcaccatcggcgtcgccgagcccgtctt 2542
                                                                                                                                                                                                                                                                                               2363 egictecgieggigagicegieggetiegecgagecigagecegagecegages 2422
                                                                                                                                                                                                                                                                                                                                                                                  2423 gctcgccatcgccgtcgccgtcgccgagctcgagtccatcctcgtcgccgtctccgtcgc 2482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2543 cgtcaccgagcccgtcttcgtcaccgtcttcgtcgccgagcccaacgccgtcgtcgtcgc 2602
                                                                                                                                                                                                                                                                                                                           257 CATCCGCGCGGCGACACCTCGCTCGCCCCCATCGCCCCTCGGCCCCGTCCAGTGGCC 316
                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This clone is available at Shanghai Hematology Institute in Shanghai.
                                                                                                                                                                                                                                                          .
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi- Tech Park, Pudong,
Location/Qualifiers
                                                                                                                                                                                                                Length 523;
                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                            Query Match 2.8%; Score 81.2; DB 10; Best Local Similarity 57.5%; Pred. No. 0.005; Matches 146; Conservative 0; Mismatches 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NPDAZH01"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /clone_lib="NPD"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2603 cggtgtcgggtggg 2616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 497 CGTGGCCGGCAGCG 510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens
                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AV752548/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM
                                                                                                                            BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERÈNCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LOCUS
```

ó g ò q ŏ g ŏ g

```
AZ386138 565 bp DNA linear GSS 02-OCT-2000 1M0144P14R Mouse 10kb plasmid UUGCIM library Mus musculus genomic clone UUGCIM0144P14 R, DNA sequence.
                                                                                                                                                       ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 (bases 1 to 565)

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Waiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mus musculus
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
                                                                                                                                                                                                                                                                                                                                                                                                                      2407 agcccgagcccgagcagctcgccatcgccgtcgccgtcgccgagctcgagtccatctcg 2466
                                                                                                                                                                                                                                                                                                                 2467 tegeegteteegtegeegteaceategeegagteegteteggteteegteaceateggeg 2526
                                                                                                                                                                                                                                                            211 AGCAGCAGCCTATGCAGTCCGTCGTCGTCGTCGTCGTCGTC-AGTCCGTCGTCGTCGTCG 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ε.,
                                                                                                                                                                                                                                                                                                                                                                     1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               s. 2030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /clone_lib="Mouse 10kb plasmid UUGC1M library"
                                                                                                  Length 1315;
                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20
                                                                                                  Score 81; DB 9; Le
Pred. No. 0.0055;
0; Mismatches 70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Laboratory Mouse DNA Resource
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0144 row: P column: 14
Seg primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
280

    .565
    /organism="Mus musculus"

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     High quality sequence stop: 565.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /db_xref="taxon:10090"
/clone="UUGC1M0144P14"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2587 acgccgtcgtcgtcgccggtgtcgg 2611
φ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /strain="C57BL/6J"
224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   32 regrestestestestes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AZ386138.1 GI:10499754
                                                                                                  Query Match 2.8%;
Best Local Similarity 65.4%;
Matches 134; Conservative
403 c
```

/tissue_type="leaf"

```
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AZ297141/c
                                             BASE COUNT
                                                                                                                                                                                                                                                                                                                                                2483
                                                                                                                                                                                                                                                                                                                                                                                                                       2543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                              ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polymucleotide kinase. Adaptor oligonicleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gelectrophoresis. Vector DNA was prepared from a derivative of pwD42 (gil4732114 gblAR129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.

1 (bases 1 to 873)
Anderson, J.M., Williams, C.E. and Goodwin, S.B.
Analysis of an EST database reveals a probable CF2 resistance gene Unpublished (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      873 bp mRNA linear EST 03-JUL-2000
EST0425 Triticum aestivum Lambda Zap Triticum aestivum cDNA clone
BE217030
                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2401 agcccgagcccgagcccgagcagctcgccatcgccgtcgccgtcgccgagctcgagtcca 2460
                                                                                                                                                                                                                                                                                                                                                                                                                                                            369 ATCCTGAGCCTGGAAAGCCTCATCCCTCTCTCTCCGTCTCCGTTTCCCTCTCCGTCTCCG 428
                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                              Length 565;
                                                                                                                                                                                                                                                                                                                                              Score 80; DB 12; Length 56:
Pred. No. 0.0076;
0; Mismatches 70; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /clone="JA1_SA_HO9_T3"
/clone_lib="Triticum aestivum Lambda Zap"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1150 Lilly Hall, West Lafayette, IN 47907, USA
Tel: 765-494-5565
Fax: 765-496-2926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Crop Production & Pest Control Research Unit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1. .873
/organism="Triticum aestivum"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              High quality sequence stop: 873.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /db_xref="taxon:4565"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Email: janderson@purdue.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /strain="P29"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BE217030.1 GI:8904570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Contact: Anderson, J.M
                                                                                                                                                                                                                                                                                                                                                2.8%;
                                                                                                                                                                                                                                                                                                                                                                al Similarity 63.5
122; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 bread wheat.
Triticum aestivum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2581 agcccaacgccg 2592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         549 TCTCCTGCCTG 560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Seg primer: T3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                USDA-ARS
                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                    Best Local
Matches 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 10
BE217030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
                                                                                                                                                                                                                                                                     BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SOURCE
```

1

ò a οp ò В ò g

ö

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              l (bases 1 to 527)

Lado, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret, A., Levins, M., Mcgann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P. and Fraser, C.M.

Mouse BAC End Sequences from Library RPCI-23

Unpublished (1999)

Other_GSSs: RPCI-23-164L13.TJ

Contact: Shaying Labo
Department of Eukaryotic Genomics

The Institute for Genomic Research
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
Library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 164 row: L column: 13
Seq primer: T7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           cgtcaccatcgccgagtccgtctcggtctccgtcaccatcggcgtcgccgagcccgtctt 2542
                                                                                                                                                                                                                                                                                                                                                      2423 gotogocatogocgtogocgtogocgagotogagtocatoctogtogocgtotocgtogo 2482
                                                                                                                                                                                                                                               2363 cgiciccgicggigagiccgicggcitcgccgagccigagcccgagcccgagcccgagca 2422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               337 CGTCCCCGCCTCCGCCGAGCCGCCCCCGAGCCGCTCCACTACGGCTCTCACGCTCTC 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    397 CGCCCCGACCGCGCCCTCGACGCCCTGCTCGCCCTACACCAGGTTCTCCGCCCACTC 456
                                                                                                                                                                                                                                                                                                    217 CATCCGCGGCGGCGACACCTCCCTGGCCGCCATCGCCGCCCTCGGCCCCGTCCAGTGGCC 276
                                                                                                                                                                                                                                                                                                                                                                                     cgtcaccgagccgtcttcgtcaccgtcttcgtcgccgagcccaacgccgtcgtcgtcgtcg
                                                                                                                                                                                             ö
                                                                                                                                     Length 873;
                                 1 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
                                                                                                                                  Score 79.6; DB 9;
Pred. No. 0.0088;
0; Mismatches 109;
/dev_stage="9 day old seedlings"
344 c 276 g 133 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /organism="Mus musculus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /strain="c57Bir/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-164L13"
/clone_lib="RPCI-23"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AZ297141.1 GI:9538926
                                                                                                                                           2.8%;
                                                                                                                                           Query Match 2.8'
Best Local Similarity 57.1'
Matches 145; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2603 cggtgtcgggtggg 2616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              457 CGTGGCCGGCAGCG 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ends
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             house mouse.
Mus musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Class: BAC
                                    119 a
```

ŏ

ŏ 임 ò g ò g

```
bread wheat.
                                                                                                                                                          ๙
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BE499920
                                                                                                                                                     .285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM
                                                                                                                                                     BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 1
BE499920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                JOURNAL
    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                    /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
ECORI, Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
ECORI sites. The ligation products were transformed into
BH108 electrocompetent cells (BRL Life Technologies). "
113 c 113 c 85 f
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CNSO53F1 1057 bp DNA linear GSS 26-JUL-2000 Tetraodon nigroviridis genome survey sequence T3 end of clone 039L09 of library A from Tetraodon nigroviridis, genomic survey
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Submitted (12-APR-2000) to the EMBL/GenBank/DDBJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Crollius, H.R., Jailion, O., Dasilva, C., Ozouf-Costaz, C., Fizames, C., Stscher, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J. Characterization and repeat analysis of the compact genome of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Bouneau, L., Fisher, C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2461 tcctcgtcgccgtctccgtcgccgtcaccatcgccgagtccgtctcggtctccgtcacca 2520
                                                                                                                                                                                                                                                                                                                                                                             2401 agcccgagcccgagcccgagcagctcgccatcgccgtcgccgtcgccgagctcgagtcca 2460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Tetraodontidae; Tetraodon.

1 (bases 1 to 1057)

Roest Crollius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W. and Weissenbach, J.

Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence

Nat. Genet. 25 (2), 235-238 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                              512 ATCATGAGCCTGGAAAGCCTCATCCCTCTCTCCTCCGTCTCCGTTTCCCTCTCCGTCTCCG 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
0
                                                                                                                                                                                                                                                                                           Length 527;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             freshwater pufferfish tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
                                                                                                                                                                                                                                                                                      Score 78.4; DB 12;
Pred. No. 0.013;
0; Mismatches 71;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GSS; genome survey sequence.
Tetraodon nigroviridis.
Tetraodon nigroviridis
/sex="Female"
/lab_host="DH10B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AL319366
AL319366.1 GI:9552250
                                                                                                                                                                                                                                                                                      2.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (bases 1 to 1057)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (bases 1 to 1057)
                                                                                                                                                                                                                                                                                                             Local Similarity 63.09
Les 121; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Direct Submission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      332 TCTCCTGCCCTG 321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2581 agcccaacgccg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Genoscope
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sednence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        20296633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20359837
                                                                                                                                                                                                 146
                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CNS053F1/c
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                               BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     JOURNAL
MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
```

```
BE499920 401 bp mRNA linear EST 04-AUG-2000 WHE0976_F08_K16ZS Wheat pre-anthesis spike cDNA library Triticum aestivum cDNA clone WHE0976_F08_K16, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Lillopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 401)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               greenhouse. Whole spike with awns trimmed, white, green and yellow anther were collected and total RNA, and poly(A) RNA were prepared, a cDNA library was made, and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 The structure and function of the expressed portion of the wheat genomes – Pre-anthesis spike cDNA library
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2486 caccatcgccgagtccgtctccgtcaccatcggcgtcgccgagcccgtcttcgt 2545
                                                                                                                                                           ID : COAAO39CF05Al~end : T3"
t 22 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /dev_stage="Adult plant"
/lab_host="Se_coll $SDRR"
/note="Vector: Lambda Uni=ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1030 CTCCGTCTCTGTCTCCCGTCTCCCTCCGTCTCCGTCTCCGTCTCCGTCTCTGTCTCCGT 971
                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20 Seq primer: Stratagene SK primer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J. Seaton, C.L. and Tong, J.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /clone_lib="Wheat pre-anthesis spike cDNA library"
/tissue_type="Spike before anthesis"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2426 egecategeegtegeegtegeegagetegagteeateetegtegeegteteegtegeegt
                                                                                                                                                                                                                                                                                                                                                            ;
0
                                                                                                                                                                                                                                                                                                DB 12; Length 1057; 0.043;
                                                                                                                                                                                                                                                                                                                                                      50; Indels
                                          /db_xref="taxon:99883"
/clone="039L09"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2546 caccgagcccgtcttcgtcaccgtcttcgtcgccg 2580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WWE0970_F08_K16"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               910 CTCTGCCTCTGTCTCCGTCTCCGTCTCAG 876
                                                                                                                                                                                                                                                                                                   Score 75; DB 1
Pred. No. 0.043
0; Mismatches
                                                                                                              /clone_lib="A"
/clone_lib="A"
/note="Genoscope sequence
Location/Qualifiers
1. .1057
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Email: oandersn@pw.usda.gov
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BE499920.1 GI:9698537
                                                                                                                                                                                                                                                                                                      2.68;
                                                                                                                                                                                                                                                                                                      Query Match 2.6'
Best Local Similarity 67.7'
Matches 105; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Triticum aestivum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unpublished (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Fax: 5105595818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Tel: 510559577
```

```
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This ender the bessen in the following URL
200-016-c08&t3-2000-02-12&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 32
High quality sequence stop: 87.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Dias Neto, E., García Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagal, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BG874723 321 bp mRNA linear EST 30-MAY-2001
RCI-BT0254-120200-016-c08 BT0254 Homo sapiens cDNA, mRNA sequence.
the cDNA clones were in vivo excised to give pBluescript phagemids in the TJ Close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 321)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Shotgun sequencing of the human transcriptome with ORF expressed
                                                                                                                                                                                                                                                                                                   2408 gcccgagcccgagcagctcgccatcgccgtcgccgtcgccgagctcgagtccatcctcgt 2467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2528 egecgagecegicting teacegagecegictine teacegictine gecgageceaa 2587
                                                                                                                                                                                                                                                                                                                                                                                             2468 egecgietecgiegécgicaceategecgagiecegieteggietecegieaceateggegi 2527
                                                                                                                                                                                                                                                                                                                                                                                                                       78
                                                                                                                                                                                                                                                                                                                              19 GCCCGGTCGCCGCGCGCGCGCGCGCCGTCAAGCTCGACCCGCGCCACTA
                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                          Length 401;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
                                                                                                                                                                                                        Score 74.6; DB 10; Length
Pred. No. 0.048;
0; Mismatches 84; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BT0254"
/dev_stage="Adult"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    cgccgtcgtcgccggtgtcgggtggg 2616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GTTCTCCGCCCACTCCGTGGCCGGCAGCG 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BG874723
BG874723.1 GI:14251654
                                                                                                                                                                                                        2.6%;
ilarity 59.8%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Tel: +55-11-2704922
Fax: +55-11-2707001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sequence tags
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens
                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Simpson, A.J.
                                                                                                                   44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     human.
                                                                                                                                                                                                                                                        125;
                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                  Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEFINITION
ACCESSION
                                                                                                                 BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BG874723/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JOURNAL
MEDLINE
COMMENT
                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LOCUS
                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
EDEMATYCE; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Wiridiplantae; Streptophyta; Elliopsida; Poales; Poaceae; Pooldeae; Triticeae; Secale.

1 (basea I to 535)
Anderson,O.D., Butler,E., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Gustafson,J.P., Han,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Ross,K., Seaton,C.L. and Tong,J.C.
The structure and function of the expressed portion of the wheat genomes - Anther cDNA library from rye
/note-"Organ: breast; vector: pucl8; Site_1: Smal; Site_2: Smal; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 ,716 - Ludwig Institute for Cancer Research) profiles into the DUC 18 vector. Reverse transcription of tissue markha and cDNA amplification were performed under low markna end cDNA applification were performed under low a 124 c 89 g 4 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEM 94335
WHE1253_D02_H03ZS Secale cereale anther cDNA library Secale cereale CDNA clone WHE1253_D02_H03, mRNA sequence.
BE494335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Uniquatishing terminal formula of Agriculture Research Service, Pacific US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center West Area, Western Regional Research Center Service, albany, CA 94710, USA Tel: 5105595773
Fax: 5105595773
Fax: 5105595718
Fax: 510559518
Famil: oandersn@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Sequence Sequ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /organism="Secale cereale"
/cultivar="Blanco"
/db_xxef="texnon:4550"
/db_xxef="texnon:4550"
/clone_lib="Secale cereale anther cDNA library"
/tissue_type="Anther"
/dev_stage="Anther"
/lab_host="E. coli SOLR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ó
                                                                                                                                                                                                                                                                                                                                                                                                                         Length 321;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                         Score 74.2; DB 10;
Pred. No. 0.055;
0; Mismatches 73;
                                                                                                                                                                                                                                                                                                                                                                                                                             2.6%; (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BE494335.1 GI:9660928
                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 2.69
Best Local Similarity 61.89
Matches 118; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Secale cereale
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2605 gtgtcgggtgg 2615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    105 rcgrrrcgrcg 95
                                                                                                                                                                                                                                                                          104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BE494335/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEFINITION
                                                                                                                                                                                                                                                                          BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

```
Site_1: EcoRI; Site_2: XhoI; Plants were grown in the greenhouse. Authers were harvested and pooled from early meiosis to late melosis. The tissue, total RNA, and poly(A) RNA were prepared (Buller, Ross and Gustafson) at University of Missouri, Columbia. A cDNA library was made, and the CDNA clones were in vivo excised to give pBluescript pagemids in the TJ Close lab (Choi, Close, Fenton) at the University of California, Riverside. Penton) at the University of California, Riverside. In the OD Anderson lab (all other authors). The Loss, Penton and DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors). The Action of California and DNA sequencing were performed of the OD Anderson lab (all other authors). The Action of California and DNA sequencing were performed of the OD Anderson lab (all other authors). The Action of California and DNA sequencing were performed in the OD Anderson lab (all other authors). The Action of California and DNA sequencing were performed on the OD Anderson lab (all other authors). The Action of California and DNA sequencing were performed on the OD Anderson lab (all other authors). The Action of California and DNA sequencing were performed on the OD Anderson lab (all other authors). The Action of California and DNA sequencing were performed on the OD Anderson lab (all other authors). The Action of California and DNA sequencing was and DNA sequencing were performed on the OD Anderson lab (all other authors). The OD Anderson lab (all other authors) and OD Anderson lab (all other authors). The OD Anderson lab (all other authors) and OD Anderson lab (all other
```

Search completed: July 2, 2002, 09:48:10 Job time: 2197 sec

2393 cgagcctgagcccgagcccgagcccgagctcgccatcgccgtcgccgtcgccgagct 2452

ŏ

Oy Oy Oy

THIS PAGE BLANK (USPTO)

```
2, 2002, 09:15:39 ; Search time 101.03 Seconds (without alignments) 813.566 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       / SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
/ SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
/ SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
/ SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
/ SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
/ SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
/ SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
/ SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
/ SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
/ SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
/ SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
/ SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
/ SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
/ SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
/ SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
/ SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*/SIDSI/gcgdata/hold-geneseq/geneseq-embl/AA1983.DAT:*/SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*/SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*/SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*/SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*/SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*/SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*/SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*/SIDSI/gcgdata/hold-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy-geneseqy-embl/AA1988.DAT:*/SIDSI/gcgdata/hold-geneseqy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 ATTQPYTWSNVAIGGGGFVD......YIGTNGRGIVYGDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /SIDSi/gcgdata/hold-geneseq/geneseqp-embl/AA1980.
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      747574
                                                         Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Total number of hits satisfying chosen parameters:
GenCore version 4.5
Copyright (c) 1993 - 2000 Comp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        747574 seqs, 111073796 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                      OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A_Geneseq_032802:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-917-376-3
                                                                                                                                                                                                                                                                                                                              July
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Database:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Searched:
                                                                                                                                                                                                                                                                                                                              Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Description	Bacterial amylase	Amino acid sequenc	Flavobacterium odo	Truncated cellulas	Polypeptide with i	Cenarchaeum symbio	Isoamylase amino a	Novel human secret	Primary transcript	Novel human diagno	Human ORFX ORF709
SUMMARIES	AAR14948	AAW59912	AAW35390	AAY13493	AAP90615	AAY90913	AAY78516	AAU31850	AAP94419	ABG08332	AAB40945
DB	12	19	19	20	10	21	21	22	10	22	21
% Query Match Length DB	1684	1291	774	1751	745	3472	776	2367	750	1045	1532
% Query Match	4.0	4.0	3.9	3.8	3.8	3.8	3.8	3.7	3.7	3.7	3.7
Score	163	159.5	158.5	152.5	152	152	151.5	150.5	148.5	148	148
Result	1	7	m	4	2	9	7	80	σ	10	11

Glutamic acid rich Drosophila melanog Human bone marrow Floculation prote S. cerevisiae FLOI Agarase 0107 from	or o	hila hila niba rymo rymo		Human potypeptide Amino acid sequenc Escherichia coli p Bankia gouldi endo Drosophila melanog Human wound healin Human potypeptide Soluble chitinase.
AAB72314 ABB59423 AAU71956 AAR47578 AAR58754 AAR60812	AAG92740 AAB35987 AAY58914 AAY68850 ABB60540 AAU27790	ABB61734 ABB63174 AAU56103 AAY34538 AAY34537 AAY34404	AAV69488 AAY58916 AAY68852 ABB71484 ABB65262	AAV13538 AAY93606 ABB52677 ABB70775 ABB40566 AAB97070 AAW02159 AAX52307
22 22 22 15 15 15	22 21 22 23 25 25	200222	5555555	22 22 22 23 21 21 21 21 21
1749 1472 1532 894 894	925 754 1000 1403 3118	1039 2464 911 1225 1226 1251	2628 1205 1205 1805 1010	230 1429 1778 875 1357 1712 2058 866
~~~~~~~~~ ~~~~~~~~~~~~~~~~~~~~~~~~~~~~				, , , , , , , , , , , , , , , , , , ,
146 144.5 143.5 143 143 143	142 140.5 140.5 140.5 139	137.5 137.5 136.5 136.5 136.5		133.5 133.5 133.5 133 133.5 133.5 133.5 132.5
12 13 14 15	18 19 20 22 23	4 2 2 2 2 2 2 4 4 4 7 8 9 C	3 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	2 6 6 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7

## ALIGNMENTS

RESULT

	tein; 1684 AA.		try)	0.	ch hydrolysis.			08669.	17595.	OCHEM IND.	, Bock A;		New bacterial amylase, A-180 for malto:pentose prodn by hydrolysis of starch, providing high yield and modifiable for secretion from host cells	pp; German.	The amino acid secuence is that of bacterial amylase A-180 whic
	AAR14948 standard; Protein; 1684 AA.	8;	1992 (first entry)	al amylase A-180.	Maltopentose; G5; starch hydrolysis	5-A.	1991.	1991; 91EP-0108669.	1990; 90DE-4017595	(CONE ) CONSORT ELEKTROCHEM IND.	Candussio A,	WPI; 1991-355676/49. N-PSDB; AAQ13939.	New bacterial amylase, A-1 hydrolysis of starch, prov secretion from host cells	Disclosure; Page 7; 21pp; German.	no acid segmence is
AAR14948	ID AAR14946	AC AAR14948;	DT 25-FEB-1992	XX DE Bacterial		AA PN EP459385-A	PD 04-DEC-1991	XX PF 28-MAY-1991;	AX PR 31-MAY-1990;	PA (CONE )	PI Schmid G,				

The amino acid sequence is that of bacterial amylase A-180 which is

~

```
47;
used to prepare maltopentose (G5) in high yields by hydrolysis of starch. It can be modified to ensure its excretion, obviating the need to concentrate and purify the enzyme, i.e. the culture supernatant can be used directly for G5 prodn. G5 yields of over 90% are possible, eliminating the need for further G5 purification. G5 is used in medical and for converting (by esterilication) fatty acids to water-soluble form, suitable for use in stable infusion solns.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          969 vttqekpegnlvtiyykkgfdtpymhyrpeggewtivpgirmeeselagyskltvdirea 1028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sklevafnngrgawdsdgennylfepgvhtylpshegrgeiipgkpgapidgnkvtiyyg 1088
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1138 --finngrg-----lwdsdnenn---yffnigdntyipgkngsageiyggkprpplvgn 1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1469 qyvdvtdiidpana-------gssagttisqmdmpqti--aidtipeggat--- 1510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----KFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDT 325.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----PKLG-W------MDEA-----MAIDPFNSDRMLY----G 398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         450 DLGGFTH-----ADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAG------ 492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QGV-----PANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVM 616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGAT - - WSQMTNFPDVGTYIANPTDTTG 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---AS 563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               YTNSWDPNDG-----AILRSSDQGA-----TWQITPLPFKLGG--NMPG 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               812 vtfewdqsdgyvveyeilrdedvvastirttftdedlnpdttytysvvavgeggqksaps 871
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FHAVPGKEG--DLW------LAASSGLYHS------TNGGSSWSAITGVS 652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----NTNKVWAAVGM 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ::|: || :| :| |: || |: || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- || :- |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       872 eal-kvttleendepkepaeapenlriaditdttvtinwnasngy-----vtg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----SFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGS--RFVW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1246 rga-----wdsdnernylfvvgnntyepgingapgqvkhgvlpddgedp----g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---IATSNTGGPYDGSSGDVW--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1186 evilyykngfdtpyvhyrpeggtwtnapgikmdkselagyskitldigradrvevafndg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     399 TGATLYATNDLTKWDSG------GQIHIAPMVKGLEETAVNDLISPPSGAPLISALG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           326 IIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWL-----TFGVOPNPPVPS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1291 diedpdhtspskptdltaiatahtvslswt--asaddvevagykiyrdgveigvtestty
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels 394;
                                                                                                                                                                                                                                                                                                                                                                                                                     Length 1684;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ------NNPVFWSRDGGATWQAVPG----------
                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
4.0%; Score 163; DB 12;
Best Local Similarity 18.7%; Pred. No. 0.022;
Matches 177; Conservative 98; Mismatches 276;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          542 AP-----GDPGQPVVYAVG------FGNSWA---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          47 ANGRWIP-----LLDWVGWNNWGYNGVVSIAADPI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --APTGFIPHKGVF--DPVNHVLY----
                                                                                                                                                                                                                                                                                                                 1684 AA;
                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1089
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      564
       8888888888888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ωp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
50;
gsnktvslyvngakirqttlpslpnwdsw-----ssktetlnln---agsntiaykydp 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This is the amino acid sequence of the mutanase enzyme, comprising the a PT box, which decomposes the alpha-1,3 glucoside bond of mutan. The mutanase enzyme is used in the method of the invention for prevention and removal of plaque and bacteria on teeth.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181 yqaesaalsggakvntdhagyigtgfvdg-ywtqgatttfsvnaptagnydvtlrygnat 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYI---ANPTD 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ptpppggniai-----gksisasshtqtyvaenandndvntywegggnpst 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            329; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  YTWSNVAIGGG-----RWDAAN 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              - used for prevention and removal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ---GATWQITPLPFKLGGNM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   G------RWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDP
                                                                                                                                                                                                           box; alpha-1,3 glucoside bond; mutan; plaque;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 1291;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                               697
                                 653 SAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
4.0%; Score 159.5; DB 19;
Best Local Similarity 19.8%; Pred. No. 0.026;
Matches 188; Conservative 105; Mismatches 329;
                                                                                                                                                                                    Amino acid sequence of the mutanase enzyme.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Disclosure; Pages 11-14; 15pp; Japanese.
                                                                                                                Ā
                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                Protein; 1291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gene encoding a mutanase enzyme of plaque and bacteria on teeth
                                                                                                                                                                                                                                                                                                      /note= "PT box"
                                                                                                                                                                                                                                                                                                                                                                          97JP-0284362
                                                                                                                                                                                                                                                                                                                                                                                               96JP-0314057
                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                 1..31
306..354
                                                                                                                                                                                                           Mutanase enzyme; PT
                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1998-474495/41
                                                                                                                standard;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1291 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                       CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NDGAILRSSDQ--
                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-PSDB; AAV53575.
                                                                                                                                                                                                                        teeth.
                                                                                                                                                                                                                                                                                                                            JP10201483-A.
                                                                                                                                                                                                                                                                                                                                                                                                                      (LIOY ) LION
                                                                                                                                                                                                                                                                                                                                                                          01-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                               25-NOV-1996;
                                                                                                                                                             20-NOV-1998
                                                                                                                                                                                                                                                                                                                                                    04-AUG-1998
                                                                                                                                                                                                                        bacteria;
                                                                                                               AAW59912
                                                                                                                                       AAW59912;
                                                                                                                                                                                                                                            Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Seguence
                                                                                                                                                                                                                                                                               Peptide
Domain
                                                                                         ~
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    350
                                                                                                    AAW59912
                                                                                                                                                                                                                                                                     Key
                                                                                         RESULT
         q
                               ò
                                                      q
                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οy
```

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
쉽
                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                             558 --NSWAASQG-VPANAQ----IRSDRVNPKTFYALSNGTFYR-----STDGGVTF- 600
       TT---GYQSDIQGVV------WVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGA 228
                                                                                 SVTSGTWTRISPVPSTDTANDYFGYSGLTID----RQHPNTIMVATQISWWPDTIIFRSTD 332
                                                                                                                                 590 snnsytnpaslvvapvsssdlvgtvswtpstpiannavsfnvnlknggtiasaggshgvt 649
                                                                                                                                                                                                                    650 vvlknasgstvqtfsgsytgslapgasvnitlpgtwtaaagsytvtatvaadanelpikg 709
                                                                                                                                                                                                                                        ----STIFTSPVFTTGTSVDYAELNP-----SIIVRAGSFD----PSSQPNDRHVAF 507
                                                                                                                                                                                                                                                                             557
                                                                                                                                                                                                                                                                                              770 psngsylqwtvrqgqggagytmrftmpdsadgm-----glngsldvyvngtkvktvslt 823
                                                                                                                                                                                                                                                                                                                                                       --QPVAAGL--PSSGAVGVMFHAVP--GKE------GDLWLAASSGLYHSTN 640
                                                                                                                                                                                                                                                                                                                                                                   GGATWIRIWDWISYPNRSLRYVLDISAEPWLIFGVQPNPPVPSPKLGWMDEAMAIDPFNS
                                                                                                                                                                                                  422 PMVKGLEETAVNDLISPPSG--APLISA-----LGDLGGFT-----HADVTAVP---
                                                                                                                                                                                                                                                                             STDGG-KNWF--QGSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFG-----
                                                                                                                                                                                                                                                                                                                               641 GGSSWSAITGVSSAV------NVGFGKSAPGSSYPAVFVVGTIG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                    1000 lrsrygqnavykgfmdnfgtnsvirdvwv----ehfecgfw----vgdygh 1042
                                                                                                                                                                                                                                                                                                                                                                                                                                 679 -----GVTGAYRS--DDCGT-----TWVLINDDQHQYGNWGQAITGDHAN 716
                                            TWQAVPGA-PTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKF-
                                                                                                                                                           393 DRMLYGTGATL----YATNDL---TKW-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1. 33
/label= sig_peptide
34. 774
/label= mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Flavobacterium odoratum isoamylase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAW35390 standard; Protein; 774
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Isomylase; industrial enzyme
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Flavobacterium odoratum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JP08023981-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        09-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              30-JAN-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAW35390;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Peptide
                                                                                                                                                                                                                                      464
                                            229
                                                                                                                      333
                                                                                                  q
                                                                                                                                        셤
                                                                                                                                                                             g
                                                                                                                                                                                                  ö
                                                                                                                                                                                                                    음
                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                 Ω
                                                                                                                      ò
                                                                                                                                                            ð
                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                             ò
```

```
44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----gfisdvdaagnrfnpnkllsdpyalelshdpttatm------tngs 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               220 VFWSRDGGATWQAV---PGAPTGFI------PHKGVFDPV---NHVLYIATSNTG 262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               288 fapdrryaydktpggptrefkemvkafhdngikvlvd-----vvynhtgeggawspt- 339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              p---tstlsgqglg------gtlyygyrawgpnw---pynaswtkgssl----- 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ANPIDITGYQSDIQG------VVWVAFDKSSSSLGQASKTIFVGVADPNNP 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         263 -----GPYDGSSGDVWKFSVTSGTWTRISPV------PSTDTANDYFGYSGLTI 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             228 itaayrgtykgaglkaaalaalgvtaieflpvqetqndtndndpsstsgdnywgymtlny 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                386 AIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEETAVNDLIS--PP---- 439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ggsgvdliaepwaiggnsyqvggfpsgwaewngayrdvvrqaqnklgsvaittg----- 499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                500 ----qmatrfagssd-----lygddgrkpwhsvnfitahdgftlkdlyscnsknn 545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              :||
600 lrsmncnnnpynldssanwlnwsrttdqnnfqsfskamiafrkahpalrpanfyssvdnn 659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----VGMYTNSWDPNDGAILRSSDQGATWQITPL 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        115 PFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYI 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DWTSYPNRSLR-----YVLDISAE-PWLTFGVQPNPPVPS------PKLGWMDEAM 385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --SGAPLIS---ALG----DLGGF-----THADVTAVPSTIFTSPVFTTGTSVDYA 481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ELNPSIIVR-AGSFDPSSQPNDRHVAFSTDGGKNWFQG---SEPGGVTTGGTVAASADGS 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           538 RFVW--APGDPGQPVVYAVGFGNSW----AASQ------GVP----ANAQ 571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ------DRVNPKTFY------ALSNGTFYRSTD-G 596
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             35 inpnklgaaydatkanvtfkvysskatrielylystatgsaekakyvmtnsggiwsvt-i 93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches 158; Conservative 103; Mismatches 265; Indels 313;
                                                                                                                                                                                                                                                          Flavobacterium odoratum isoamylase gene – useful in industrial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 774;
                                                                                                                                                                                                                                                                                                                                                                                                  The present sequence is Flavobacterium odoratum isomylase, which can be used as an industrial enzyme.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 3.9%; Score 158.5; DB 19;
Best Local Similarity 18.8%; Pred. No. 0.016;
                                                                                                                                                                                                                                                                                                                                          Claim 1; Page 6; 8pp; Japanese.
                                                                                                                (NISO ) NIPPON SHOKUHIN KAKO
94JP-0167267
                                                        94JP-0167267
                                                                                                                                                                     WPI; 1998-021932/03.
N-PSDB; AAT95400.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           572 IRS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        76 INTNKVWAA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     774 AA;
     20-JUL-1994;
                                                        20-JUL-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                        processes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                94
```

```
The invention relates to a recombinant cellulase active protein free of proteinases of native thermophilic and alkalinophilic origin, comprising the truncated sequences Cel B5, Cel B4/5, Cel B1, Cel B12/3, Cel 12/3, Cel 6 or Cel 59765, or a stability region from one of the defined full-length sequences, or functional equivalents. Cel B5 extends from amino acid A1011 to P1424 or K1425 or N1426, and Cel B4/5 extends from amino acid K635 to N1426 in the sequence shown in A4713492, Cel E1 extends from amino acid X99 to D481, Cel E6 extends from maino acid V1233 to K1751 and the stability region extends from maino acid V1233 to the sequence shown in A4713493, Cel E1/2 extends from maino acid V1233 to the sequence shown in A4713493, Cel E1/2 shown in A4713494. The new enzymes are useful in laundry detergent compositions to prevent or remove staining, backstaining or graying, for use on cellulosic materials including cotton-containing fabrics. They are especially useful processing of textiles where cellulose breakdown is required. The new processing of textiles where cellulose breakdown is required. The new contained enzymes show reduced redeposition of dye compared to using
597 GVTFQPVAAGLPSSGAVGVMF-----HAVPGK-EGDLWLAASSGLYHSTNGGSSWSAIT 649
                                                                            gnvmeqlrwfkpdggvadatyfndannhaiawridgsefgdtasaiyvahn---awsa-- 714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Cellulase; proteinase; truncated; Cel B5; Cel B4/5; Cel E1, Cel E1/2; Cel 1/2/3; Cel 6; Cel E3/B5; laundry detergent; stain-removing; cotton-containing fabric; stonewashing.
                                                                                                                                                    650 GVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQ 708
                                                                                                                                                                                                                                 -----gvnftlpwpgagkswyrvtdtcgwaegasgvgapgse-alvggentayglcgr 766
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New truncated cellulase proteins, useful in detergents and for producing 'stonewashed' denim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Farrington GK;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Truncated cellulases comprising amino acid sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bergguist PL, Daniels RM, vrgan H, Williams DP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         non-truncated cellulase compositions
                                                                                                                                                                                                                                                                                                                                                                                                                              AAY13493 standard; Protein; 1751 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 7; Page 37-41; 65pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (CLRN ) CLARIANT FINANCE BVI LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98EP-0810919.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  97US-0932571.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Anderson P, Berya-
- Mn, Morgan H,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1999-315403/27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1751 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-PSDB; AAX55662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19-SEP-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          15-SEP-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30-JUL-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EP921188-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             09-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY13493;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                            099
                                                                                                                                                                                                                                 715
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAMES OF COLOR COL
                                                                        g
                                                                                                                                                        ò
```

```
360 EPWLTFGVQPNPPVPSPKLGW----MDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSG 415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                nvtfnfvklssgvsgady-ylevgfssgagqlqpgkdtgdiqvrfnkndwsnyngaddws 799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             916
                                                                                                                                                                                                                                                                                                                                                                                  534
                                                                                                                                                                                                                                                                                                                                                                                                                      SGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVL-DISA 359
                                                                                                                                                                                                                                                                                                                                                                                                                                                       535 s-----pdvvkvdt-----yyieggkisgpyvwd----knrniyyvlvdfsg 572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               573 tk----iypggevehkkgaqfkisvpggypwdptn-dpsykgltsgleknkyiaaydnn 626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        n----lvwglepgaatstpaptstptptptptvtatptptptptptptptgspgtgsgyk 680
                                                                                                                                                                                                                                       ------ahsswansmkipeyhrhilygalvggpgsddsyndditdyvgne 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               860 vsatptpaptaspvg-gsywtpsesygalkvwyangnlssptnvlnpk--ikienvgtta
                                                     46 AANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVCM-YTNSWDPNDGAILRSSD
                                                                                                                                                                                                                                                                         193 VAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNH
                                                                                                                                                                                                                                                                                                             681 vlyknnetsastgsirpwfkivnggsssvdlsrvkirywytvdgdkpgsavcdwaqigas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           917 vdlsrvkvrywytidgeatgsvsvassinpayidvkfvklganaggadyyveigfksgag
                                                                                         290 vrygtlimlakitgkel--ykgavernldhwtdritytpkgmayltgw----gslryatt
                                                                                                                                                                                                                                                                                                                                                VLYIAT -- SNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTAN-DYF------GY
                                                                                                                                                                                                                                                                                                                                                                       ::: : |: || : :| : || : ||
eifveskfgnsqgp---nytevisylynrtgw----pprvtdklsfkyfidlteliqagy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSA-----ITGVSSAVNVGFGKSAPGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ------FYRSTDGGVT------FQP-----VAAGLPSSGA------VG
                                                                                                                            105 QG-----ATWQ-----ITPLPFKLGGNMPGR----GMGERLAVDPNNDNI
                                                                                                                                                      LYFGAPSGKGLWRSTDSGATWSQMTNFPD-----VGTYIANPTDTTGYQSDIQGVVW--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----GOPVVYAVGFGNSWAASQGVPA------NAQIRSDRVNPKTFYALSN-GT--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---PSGAP---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STIFTSPVFTTGTS-VDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKN----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --GSEPGGVTTGGTVAASADGSRFVWAPGDP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----LISALGDL------GGFTHADVTAV------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SYPAVFVVGTIGGVTGA----YRSDDCGTT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAP90615
ID AAP90615 standard; protein; 745 AA.
XX
AC AAP90615;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    416 GQIHIAPMVKGLEETAVNDLISP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               712 GDHANLRRVYIGTNGRGI-VYGDIG 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          515 WFQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         S
                                                                                                                                                                                                    141
                                                                                                                                                                                                                                                                                                                                                                                                                      301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                547
                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                              à
                                                                                                                                                        g
                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                            pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

49;

Gaps

YTWSNVAIGGGGFVDGIVFNEGAPGILYVRTD------IGGMYRW-----D 45

Match 3.8%; Score 152.5; DB 20; Length 1751; Local Similarity 18.7%; Pred. No. 0.12; les 184; Conservative 101; Mismatches 349; Indels 351;

Query Match

Matches

ò

```
DRMLY -----GTGATLY ---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Cenarchaeum symbiosum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Feldman
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (DIVE-) DIVERSA CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2000-293148/25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-PSDB; AAA55188.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200018909-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RV,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              29-SEP-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     29-SEP-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30-AUG-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        06-APR-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY90913;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Swanson
  393
                                           411
                                                                                                                        469
                                                                                                                                                                                                                                                                                                                                                                                                                                                          999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          707
                                                                                 436
                                                                                                                                                                                                                                                  545
                                                                                                                                                                 491
                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                 ò
                                                                                                                    q
                                                                                                                                                                 οy
                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            43;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     expressed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         malt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polypeptide (see AAN90395) with isoamylase activity. The DNA is obtd. from Pseudomonas amyloderamosa, pref. strain SB-15, and expresse recombinantly. The polypeptide has mol. wt. 75-85 kD and isoelectric point 4.6-4.8. It is used, together with beta-amylase, to convert amylaceous materials to high maltose prods. from which high purity maltose can be recovered. See AAP90610, AAN90390 for signal peptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      370 gfr----fdlasvlgnsclnavhasa------ 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                51 WIPLLDWVGWNWGYN---GVVS---IAADPINTNKVWAAVGMYTNSWDPNDGAILRSSD 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             105 QGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQM 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---QASKTIFVGVA------DPNNPV--------FWSRDGGATWQA 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                            Polypeptide with isoamylase activity from Pseudomonas amyloderamosa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----gnvfasahyrttdsgi----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           : | : | : | : | : | : | : | : | : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   aglkasylaslgvtaveflpvqetqndandvvpnsdanqnywgymtenyfspdrryaynk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   233 VPGAPTG-FIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polypeptide with isoamylase activity for converting starch t
derived from Pseudomonas amyloderamosa or transformed host
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels 292;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       82 wgp--nwpyasnwgkgsgagcvsdvdangdrfnpnkllldpyagevsgdp----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               165 TNFPDVGTYIANPTDTTG------YQSDIQGVVWVAFDKSSSLG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; Score 152; DB 10; Length 745;
; Pred. No. 0.042;
91; Mismatches 273; Indels 29
                                                                                                     amyloderamosa; isoamylase; starch; malt;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 5; page 22; 29pp; English.
                                                                                                                            maltose
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3.8%;
ilarity 19.6%;
Conservative 9
                                                                                                                                                                                                                                                                                         89EP-0301088
                   (first entry).
                                                                                                                                                                                                                                                                                                                                88JP-0024762
                                                                                                                                                                                                                                                                                                                                                                       (HAYB ) HAYASHIBARA SEIBUTS
                                                                                                                                                                 Pseudomonas amyloderamosa.
                                                                                                                          amylaceous materials;
                                                                                                                                                                                                                                                                                                                                                                                                               Futai M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1989-229398/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; AAN90395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             microorganisms.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      745
                                                                                                                                                                                                                                                                                     03-FEB-1989;
                                                                                                                                                                                                                                                                                                                                04-FEB-1988;
                   03-NOV-1989
                                                                                                     Pseudomonas
                                                                                                                                                                                                                                              09-AUG-1989
                                                                                                                                                                                                                                                                                                                                                                                                                  Amemura A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Sim
Matches 160;
                                                                                                                                                                                                       EP327391-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          is
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 204
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
```

ð

ò

ò

```
New nucleic acids and proteins isolated from the non-thermophilic crenarchaeote Cenarchaeum symbiosum, useful in characterizing the physiology of these archae and in therapeutic, industrial or laboratory techniques -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Cenarchaeum symbiosum open reading frame protein sequence SEQ ID NO:4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAA55186 to AAA55226 and AAY90913 to AAY90951 represent nucleic acids and proteins isolated from the non-thermophilit cranarchaedcte. Cenarchaeum symbiosum. The nucleic acids and proteins identified in the present invention are useful in characterising the physiology of these archae and can be used in therapeutic, industrial or laboratory techniques. AAA55227 to AAA55260 represent promoter sequences from Cenarchaeum symbiosum. AAA55261 to AAA55269 represent PCR primers and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ngps----lgdsnsiyvayn-----gwsssvtftlpappsgtqwyrvtdt 706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               physiology;
                                                                                                                                                                                                                                DP----GQPVVYAVGFGN------SWAASQG---VPANAQIRSDR---VNPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   571 tplmqqqdeylrtlqcnnnaynldssanwltyswttdqsnfytfaqrlirsarhiplrps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    640 NGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDD
                                                                                       nrilreftvrpaaggtvwiclrnlgpsaatr--tswvdsrrvvrvewsvprglrqaqnel
                                                                                                                                                                                                                                                                                                                                                       --AGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTV---AASADGSRF-VWAPG
                                                                                                                                                                                                                                                                                                                                                                                                                                          518 ngannsqasygpsdg----gtstnyswdggmsag---tgaavdgrraartgmafemlsag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          581 TFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASS-GLYHST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             631 swysgsqltwy-----qp----sgava-----dsnywnntsnyaiayai
----ATNDLTKW-DSGQIHIAPMVKGLEETAVNDL
                                                                                                                                                                         ISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVR----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Cenarchaeum symbiosum; non-thermophilic; crenarchaeote; physicharacterisation; archae; therapeutic; industrial; laboratory
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----cdwndgastfvapgsetlig--gagttyggcg 736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             700 QHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDIG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 26; Page 102-110; 210pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ΰ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY90913 standard; Protein; 3472 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Schleper
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98US-0102294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RA,
```

ø

```
202 LGQASKTIFVGVAD-----PN---NPVFW-----SRDGGATWQAVPGAPTG-F 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    148 agevsqdp-----lnpsnq-----ngnvfasgasyr----ttdsgiy--ap 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     96 gitgavy-----ygyrawgpnwpyasnwgkgsqagfvsdvdangdrfnpnkllldpy 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           243 lg-vtaveflpvgetgndandvvpnsdangnywgymtenyfspdrryaynkaaggptaef 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        27 GAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVG 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     87 MYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SGKGLWRSTDSGATWSQMTNFPDV -- GTYIANPTD -TTGYQSDIQGVVWVAFDKSS -- SS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                183 kgvvlvpstqstgtkptraqkddviyevhvrgfteqdtsipaqyrgtyygaglkasylas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 20.68
Matches 164; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   147
    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                           50;
                                                                                                                                                                    986 dgtvvflsd-gptgrvypyslnipfdissaapg----gfvivpvgvsdiafsa-dgrnm 1038
                                                                                                                                                                                                                                 1039 lvadetggihrylarspyeigtdfiksslgefvetfsaaprvqdlagiafshdgmimlaa 1098
                                                                                                                                                                                                                                                                                                                                                           1158 aapygi----geaeplpplflgvgaeeatlspdgrhilv----pgrpglsgyslfstnl 1208
                                                                                                                                                                                                                                                                                                                                                                                                            1328 lddaaagkisgytspvfrtwsspflgtdgatrphtlgfgdvrladiydasgd----- 1379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----vpspsgiefsddgmrm----fvtgigtpginiftlsapfditlpkhsgstnig 1428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            || :|:|| :||| || ---gl---svsdlafanngnsltvldvdgvlrvyalgddynvvtgttqkfritldttgg 1481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1482 ipnsiytspdglsqfvayddridlyvlgspndisstteiipyslprp---dpptgmdftp 1538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    : | : ||: | : : ||| ||| || dgrrmflstengidgyllsepfavttsvflrtipidggaegirfvdngrglfvpgadgii 1598
                                                                                                                                                                                                                                                                                              1099 ggsgsvhryslpspyavsgakyeetamig-gspsglefssdglrmfvpdagsetaavygl 1157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1268 megvtldsregtlrvrldrtvdvdtvrpykmwvedsdgsqttlanstllnaensnillfr 1327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1599 grheliypygastslletvrdgv----------tdggpgenpa 1631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1632 ageirlagt----fnasdn----vqspsgiefsgdgtgmf--vtgfgaagvnefslsa 1679
                                                                                                                                                                                                                                                                   235
                                                                                                                                                                                                                                                                                                                                264
                                                                                                                                                                                                    152 WRSTDSGATWSQMTNFP-DVG-----TYIANP--TDTTGYQSDIQGVVWVAF 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---IIFR 329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STDGGA------TWTRIW---DWTSYPNR----SLRYVLDISAEPWLTFGVQ 368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PNPPVPSPK-LGWMDEAMAIDPFNSDRMLYGTGA-----TLYATNDLTKWDSGGQIHIA 421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PMVKGLEETAVNDLISPPSGAPL----ISALGD----LGGFTH----ADVT-A 461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   501 NDRHVAFSTDGGKNWFQGSEPGGVTTGG----TVAASADGSRFV-----WAPGDPG-- 547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSA 663
                                                                                                                                        DGAILRSSDQGATWQITP----LPFKLGGNMPGRCMGERLAVDPNNDNILYFGAPSGKGL 151
                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       462 VPSTIFTSP-----VFTTGTSVDYA---ELNPSIIVRAGSFDPSS----OP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -- QPVVYAVGFGNSW--AASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPV
                                                                                                                                                                                                                                                                                                                             -----NHVLYIATSNTGGP------
                                                                                                                                                                                                                                                                                                                                                                                          -----YDGSSGD--VWKFSVTSGTWTRISPVPSTDTAN-----DYFGYSG-----
                                                                                                                                                                                                                                                                   DKSSS------SLGQASKTIFVGVADPNNPVFWS-----RDGGATWQAVPG-
                                                                           Query Match 3.8%; Score 152; DB 21; Length 3472; Best Local Similarity 21.0%; Pred. No. 0.33; Matches 172; Conservative 100; Mismatches 244; Indels 304;
probes used in examples from the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1680 pfdttlpvhvelhdigggpavdlafaed--grtlillaad 1717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P-GSSYPAVFVVGTIGG----VTGAYRSDDCGTTWVLINDD 699
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---LTIDRQH-----PNTIMVAT---QISWWPDT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY78516 standard; protein; 776 AA
                                                                                                                                                                                                                                                                                                                             236 -APTGFIPHKGVFDPV------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                3472 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           05-MAY-2000
                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY78516;
                                                                                                                                                                                                                                                                 196
                                                                                                                                                                                                                                                                                                                                                                                          265
                                                                                                                                                                                                                                                                                                                                                                                                                                                      303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY78516
ID AAY7
XX
AC AAY7
XX
DT 05-M
   SXS
                                                                                                                                                                     g
                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                          ò
                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                         엄
                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                        Qγ
                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
This sequence represents an isoamylase amino acid sequence. Isoamylases are debranching enzymes which hydrolyse alpha-1,6-D-glucosidic branch linkages in amylopectin. This sequence represents a parent enzyme from which genetically engineered variant debranching enzymes can be created. The genetically engineered enzyme has an improved thermostability at a pH confidence of 4-6 compared to the parent enzyme. The modified enzyme and or glycogen compared to the parent enzyme. The modified enzyme and or glycogen compared to the parent enzyme. The argin anylopectin and or glycogen compared to the parent enzyme. The starches such as corn, potato, wheat, manico and rice starch are useful when starches such as corn, potato, wheat, manico and rice starch are used as starting materials in commercial large scale production of sugars, such as high fructose syrup, high maltose syrup, maltodextrins, amylose, (4-66 oligosaccharides and other carbohydrates such as fat replacers. The chermostable debranching enzymes of the invention make it possible to perform liquefaction and debranching using the thermostable debranching carcharitication step. By debranching using the thermostable debranching enzyme variants during liquefaction together with the action of an alpha-amylase, the formation of panose precursors is reduced, thereby reducing the panose content in the final product and increasing the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Variant bacterial pullulanases and isoamylases having, e.g. increased thermostability, used for converting starch from potatoes into high fructose svrnn
                                                                                          Isoamylase; debranching enzyme; amylopectin; thermostable; corn; genetically engineered variant; wheat; potato; wheat; manioc; rice starch; sugar production; high fructose syrup; high maltose syrup; maltodextrin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Svendsen A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 29; Fig 3; 116pp; English.
Isoamylase amino acid sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99WO-DK00381.
                                                                                                                                                                                                                                                                                                                                            Pseudomonas amyloderamosa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (NOVO ) NOVO-NORDISK AS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bisgard-Frantzen H,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2000-160767/14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     776 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200001796-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               fructose syrup
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 02-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 02-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-JAN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
```

42;

Indels 213;

DB 21; Length 776;

3.8%; Score 151.5; DB 21; 20.6%; Pred. No. 0.048; iive 91; Mismatches 329;

us-09-917-376-3.rag

```
445
                                                                                                           DWTS-YPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAID----PFNSDRML 396
                                                                                                                                                                                                                        493
                                                                                                                                                                                                                                              503 yviqdandfsgs----snlfgssgrspwnsinfidvhdgmtlkdvyscngannsgawp 556
                                                                                                                                                                                                                                                                                                                                   605
                                                                                                                                                                                                                                                                                                                                                                                                                                               FGKSAPGSSYPAVFVVGTIGGYTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLR 718
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         698 -gpslgdsnsiyvayngwsssvtftlpappsgtqwyrvtdt----cdwndgastfvapgs 752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             stem cell proliferation; haematopoiesis; nerve tissue regeneration;
immune suppression; immune stimulation; anti-inflammatory; leukaemia
                           qamvqafhnagikvymdv------vynhtaeggtwtssdpttatiyswrgldn
                                                     -----DTANDYF-GYSGLTIDRQHPNTI----MVATQISWWPDTIIFRSTDGGATWTRIW
                                                                                                                            --TAVNDLISPPSGAPLI
                                                                                                                                                                                                                      SALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVR------AGS
                                                                                                                                                                                                                                                                                                        ygps------dggtstnyswdggmsag---tgaavdgrraartgmafeml--sa
                                                                                                                                                                                                                                                                                                                                   GQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTF-YALSNGTFYRSTDGGVTFQPVAA
                                                                                                                                                                                                                                                                                                                                                            gtplmgg---gdeylrtlgcnnnaynldssanwltyswttdgsnfytfagrliafrkahp
                                                                                                                                                                                                                                                                                                                                                                                        GL-PSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNG----GSSWSAIT--GVSSAVNVG
                                                                                                                                                                                                                                                                                                                                                                                                        atyyeltsgngyfydntgiganfntyntvagnlivdslaywantm---gvdgfr----f
                                                                                                                                                                                           452 ggsgldlfa----epwaiggnsyglggfpggwsewnglfrdslrgagnel----gsmti
                                                                                                                                                                                                                                                                            FDPSSQPNDRHVAFSTDGGK----NWFQGSEPGGVTTGGTV---AASADGSRFVWAPGDP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gene therapy; nutritional supplement;
241 IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPST
                                                                                                                                                                  YGTGATLYATNDLTKWDSGQIH-IAPMVKGLEE---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAU31850 standard; Protein; 2367 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Novel human secreted protein #2341.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Drmanac RT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18-APR-2000; 2000US-0552929.
26-JAN-2001; 2001US-0770160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16-APR-2001; 2001WO-US08656
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RVYIGTNGRGIVYGDIG 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            753 etlig--gagttygqcg 767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                vaccination;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2001-611725/70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (HYSE-) HYSEQ INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Liu C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200179449-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .18-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Fang YT,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAU31850;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Human;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                           302
                                                                                 349
                                                                                                           342
                                                                                                                                     401
                                                                                                                                                                                                                                                                                                        557
                                                                                                                                                                                                                                                                                                                                547
                                                                                                                                                                                                                                                                                                                                                            900
                                                                                                                                                                                                                                                                                                                                                                                        909
                                                                                                                                                                                                                                                                                                                                                                                                                   657
                                                                                                                                                                                                                                                                                                                                                                                                                                               629
                                                     292
                                                                                                                                                                 397
                                                                                                                                                                                                                      446
                                                                                                                                                                                                                                                                            464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                            g
                                                                                                                                  g
                                                                                                                                                                                         q
                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                           ò
                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
 ð
                                                     ò
                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                            à
```

```
43;
                                                                                                                 The invention relates to novel human secreted polypeptides. The polypeptides are useful for determining the presence of or predisposition to a disease associated with altered levels of polypeptide. The polypeptides are also useful for identifying agents (agonists and antagonists) that bind to them. Cells expressing the proteins are useful for identifying a therapeutic agent for use in treatment of a pathology related to aberrant expression or physiological interactions of the polypeptide. Vectors comprising the nucleic acids encoding the polypeptides and cells genetically engineered to express them are also useful for producing the proteins. The proteins are useful in genetic accentation, testing and therapy, and can be used as nutritional supplements. They may be used to
                                                                                                                                                                                                                                                                                                                                                                                                         increase stem cell proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon and/or nerve tissue growth or regeneration; immune suppression and/or stimulation; as anti-inflammatory agents; and in treatment of leukaemias. AMUS9510-AMU33304 represent the amino acid sequences of novel human secreted proteins of the invention.
Nucleic acids encoding a range of human polypeptides, useful in genetic vaccination, testing and therapy {\ }^{-}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----wiaaerwikfrfqtmdgdwensvswa 1360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----lagnqa--igelwaagqasdgswpggqasgvswvgeeaiggs--wt---gaenqas 1504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -------gagg-----qasggsmlgpe-----dgssgrswadtadgas 1591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                .623 enqtteegswagaggqagggskvgpedqssgrswansgdqi---sggflvgivdqanggs 1679
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   dqasgwfcvcpgsqtnggswggasgqdvggsr----pgptnqssagswdspgsqv---- 1784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1785 sgscwtgagavdqaggcskpgfedqai----gggfwpgagdqtggggsrpgsedqssgigs 1840
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTD-SGATWSQMTNFPDVGTYIA 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   406 TNDLTK---W-----DSGGQIHIAPMVKGLEETAVN--DLISPPSGAPLISALGDL--GG 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3 TQP----YTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWW 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   346 YPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ggsr-lghvdqssggaw-----gggskprf
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DGGKNWF---QGSEPGGVTTGGTVAASADGSRFVWAPG------DPGQPVVYAVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1403 ----nvvgigtwarageqasgg--lwaggqtsegtwagdkasggawtgaenqasggswa-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      234 PGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSS ...--GDVWKF---SVTSGTWTRI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1505 egswag--agagnmssvsywagvvdqagggswagtsdqsgggskprfedqasgegswa--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FTHADVTAVPSTIFTSPVFTTGTS--VDYAELNPSII--VRAGSFDPSSQPNDRHVAFST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GWNNWGYNGVVSIAADPINTNKVWAAVGMYTN--SWDPNDGAILRSSDQGATWQITPLPF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1361 ddeneasigswsgasdkagiirswavacdetsvksw-----agarae----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFY----RSTDGGVT--FQPVAAGLPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        176 NPTDTTGYQSDIQGVVWVAFDKSSSL--GQASKTIFVGVADPNNPVFWSRDGGATWQAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 2367;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3.7%; Score 150.5; DB 22;
20.5%; Pred. No. 0.25;
iive 84; Mismatches 310;
                                                                      20; Page 519; 765pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2367 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Sin
Matches 160;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                           Claim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1561
     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
446 SALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVR------AGS 493
                                                                                                                                                                                                                                                            ------DTANDYF-GYSGLTIDRQHPNTI----MVATQISWWPDTIIFRSTDGGATWTRIW 341
                                                                                                                  342 DWTS-YPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAID----PFNSDRML 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              659 FGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLR 718
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gamvgafhnagikvymdv------vynhtaeggtwtssdpttatiyswrgldn
                                                                           574 gtplmgg---gdeylrtlgcnnnaynldssanwltyswttdgsnfytfagrliafrkahp
                                                                                                                                                                                                                                                                                                                                                                                                                    606 GL-PSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNG----GSSWSAIT--GVSSAVNVG
                                                                                                                                               375 dlasvlgnscingaytasa-----pncpnggynfdaadsnvainrilreftvrpaa
                                                                                                                                                                             397 YGTGATLYATNDLTKWDSGGQIH-IAPMVKGLEE------TAVNDLISPPSGAPLI
                                                                                                                                                                                                                                                                                                 494 FDPSSQPNDRHVAFSTDGGK----NWFQGSEPGGVTTGGTV---AASADGSRFVWAPGDP
                                                                                                                                                                                                                                                                                                                                                          547 GQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTF-YALSNGTFYRSTDGGVTFQPVAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
241 IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPST-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New isolated polynucleotide and encoded polypeptides, useful in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Novel human diagnostic protein #8323.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABG08332 standard; Protein; 1045 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Tang YT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           31-MAR-2000; 2000US-0540217. 23-AUG-2000; 2000US-0649167.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              30-MAR-2001; 2001WO-US08631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      719 RVYIGTNGRGIVYGDIG 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13-FEB-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Drmanac RT, Liu C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2001-639362/73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (HYSE-) HYSEQ INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-PSDB; AAS72519
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WO200175067-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11-OCT-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ABG08332;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 10
                             276
                                                         292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ABG08332
                          g
                                                         ò
                                                                                  g
                                                                                                                  Qγ
                                                                                                                                              g
                                                                                                                                                                             δ
                                                                                                                                                                                                         유
                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            42;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gene isolated from Pseudomonas inserted into an expression vector and thence into a host, producing large quantities of isomylase in a suitable medium. Especially useful in producing maltose and amides.
                                      1841 wgvaggqvlggarpgpadqssggsw--aqtg---nqssgrsw-----igpgdqa 1884
                                                                                      | | :| | | | :| | :| | 1885 vdcskpefedqacgggswagagsqasgeswagsrpgneaiggsrmgsedqatggswars 1943
            SGAVG--VMFHAVPG----KEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSA 663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          87 MYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAP 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    122 aqevsqdp----lnpsnq-----ngnvfasgasyr-----ttdsgiy--ap 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LGQASKTIFVGVAD-----PN---PN---NPVFW-----SRDGGATWQAVPGAPTG-F 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               147 SGKGLWRSTDSGATWSQMTNFPDV--GTYIANPTD-TTGYQSDIQGVVWVAFDKSS--SS 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              70 gitgavy-----ygyrawgpnwpyasnwgkgsgagfvsdvdangdrfnpnkllldpy 121
                                                                    664 PGSSYPAVFVVGTIGGVTGAYRSDDCGTTWV------LINDDQHQYGNWGQA 709
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  27 GAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVG 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3.7%; Score 148.5; DB 10; Length 750; 20.6%; Pred. No. 0.073; Live 90; Mismatches 330; Indels 213;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            isolated from Pseudomonas species and cloned to produce polypeptide(s) having biological activity of isoamylase.
                                                                                                                                                                                                                                                                    Primary transcript of Pseudomonas SMP1 isoamylase gene.
                                                                                                                                                                           AAP94419 standard; protein; 750 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gene encoding isoamylase enzyme
isolated from Pseudomonas specie
                                                                                                                                                                                                                                                                                                 Isoamylase; maltose; amides; ss
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 10; Fig 9; 45pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                    88EP-0830322.
                                                                                                                                                                                                                                                                                                                                                                                                                                              87IT-0021615.
88IT-0021282.
                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 3.79
Best Local Similarity 20.69
Matches 164; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ENIE ) ENIRICERCHE SPA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1989-040968/06.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               750 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAN91089
                                                                                                                                                                                                                                                                                                                             Pseudomonas sp.
                                                                                                                                                                                                                                                                                                                                                                                                                  26-JUL-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                               07-AUG-1987;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            08-JUL-1988;
                                                                                                                                                                                                                                     19-JUN-1990
                                                                                                                                                                                                                                                                                                                                                                                      38-FEB-1989
                                                                                                                                                                                                                                                                                                                                                         EP302838-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Tognoni A;
                                                                                                                                                                                                         AAP94419;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                202
                                                                                                                                                             AAP94419
                                      9
                                                                                               원
          ŏ
                                                                       ò
                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

us-09-917-376-3.rag

qwtsasastspdtaaamththqa-esteasgqtqtsepassgsrttsagtatpsssg---

759

g

```
The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences of the invention.

Conduct amino acid sequences of the invention of mutations diagnostic amino acid sequences of the invention.

Conduct man acid sequences of the invention.

Conduct while was obtained in electronic format directly from WIPO at the vipo int/pub/published_pct_sequences.
diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess
                                                                                                          Claim 20; SEQ ID No 38691; 103pp; English.
                                                           biodiversity
```

1045 AA; Sequence

```
26;
                                                                                                                                                                                                                                                                                                                                               357 vlsspsgfnpsgtvsgetfpsgettisspssvsntflvtskvfrmpisrdstlgnteets 416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             452 GGFTHADVT------RAVPSTIFTSPVFTTGTSVDYAELNPSIIV------RAGSFD 495
                                                                                                                                                                                                                                                                                                                                                                                                         ------ 312
                                                                                                                                                                                                                                                                                                                                                                                                                                   :: ||| : |: || || : || || : || 417 lsvsgtisaitskvstiwwsdtlstalspsslppkistafhtggsegaettgrpherssf 476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----IMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTF 365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 svtqmikt-atspssspmldrhtsqqittap---stnhstihststspqespavsq---- 641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 paptgdghttqapttalqatpsshdatlgpsggtsl---sktgaltlansvvstpggpeg 758
                                                 Indels 194; Gaps
                                                                                                  145 APSGKGLWRSTDSGATWSQMTN----FPDVGTYIANPTDT----TGYQSDIQGVVWVA 194
                                                                                                                                               245 spsgetatsslcsvtntsmmtsekitvttstgstlgnpgetssvpvtgslmpvtsaalvt 304
                                                                                                                                                                                                     -----VGVADPNNPVFWSRDGGATWQA 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -- LGWMDEAMAIDPFN 391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SDRMLYGTGATLYATNDLTKWDSGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDL 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              642 rghtqapqttqesqttrsvspmtdtktvttpgssftasghspseivpqdaptisaattfa 701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   540 VWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVT 599
                                                                                                                                                                                                                                                    305 vdpegqspatfsrtstqdttafsknhqtqsvettrvsqintlntltpv···---ttst
                                                                                                                                                                                                                                                                                                   233 VPGAPTGFIPHKGVFD---PVNHVLYIATSNTGGPY-------DGSSGDVWKFS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | ::||:|||:|||
477 spgvsqeiftlhetttwpssf---sskghttwsq----telpststgaatrlvtgnpstg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           530 aagtiprvpskvsaigepgepttysshsttlpkttgagaqtqwtqetgttgeallsspsy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PS-------SQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRF
  DB 22; Length 1045;
3.7%; Score 148; DB 22;
19.3%; Pred. No. 0.12;
iive 83; Mismatches 280;
                                                                                                                                                                                                                                                                                                                                                                                                    277 VT-SGTWTRISPVPST----DTANDYFGYSGL---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -
--
                                                                                                                                                                                                   195 FDKSSSSLGQASKTIF-----
                       Best Local Similarity 19.3
Matches 133; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         366 GVQPNPPVPS-
  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       313
                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ô
```

```
AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397, which represent the human ORFX open reading frames 1 to 3161. The ORFX sequences have activities such as: cytostatic; hepatotropic; vulnerary; antipsorlatic; antiparkinsonian; nootropic; neuroprotective; osteopathic; anticonvulsant; thrombolytic; immunosuppressant; antidinabetic; hypotensive; dermatoblytic; coagulant; vasofant; thrombolytic; antidinabetic; hypotensive; dermatological; immunosuppressive; antidinflammatory; antibacterial; antiviral; antifungal; antirheumatic; antithyroid; and antiansemic. The sequences can be used for determining the presence of or predisposition to, or preventing or treating pathological conditions associated with an ORFX-associated disorder. The
                                                                                                                                                                                                                                                                                                                                                                        Human; open reading frame; ORFX; detection; cytostatic; hepatotropic; vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective; anticonvolsant; osteopathic; antiarthritic; immunosuppressant; cardiant; immunostimulant; thrombolytic; coaquiant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antinflammatory; antianaemic; gene therapy; cancer; proliferative antithyroid; antianaemic; gene therapy; cancer; proliferative disorder; hypertension; neurodegenerative disorder; osteoarthritis; graft vs host disease; cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; cholesterol ester storage; systemic lupus erythematosus; infection; severe combined immunodeficiency; malaria; autoimmune disorder; asthma; allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; antiinflammatory disease; coaquiation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             frame X,
Novel nucleic acids and peptides derived from open reading useful for treating e.g. cancers, proliferative disorders, neurodegenerative disorders and cardiovascular disease
                                                                                                                                                                                                                                                                                                                                      Human ORFX ORF709 polypeptide sequence SEQ ID NO:1418.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 11; Page 1203-1206; 5507pp; English.
                                                                           653 SAVNVGFGKSAPGSSYPAVFVVGTIGGVTG 682
                                                                                                               870 ssi-----vpgtfhptlseastagrptg 892
                                                                                                                                                                                                                   AAB40945 standard; Protein; 1532 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0127607.
99US-0127636.
99US-0127728.
2000US-0540763.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       31-MAR-2000; 2000WO-US08621
                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              thrombosis; contraceptive.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Shimkets RA, Leach M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (CURA-) CURAGEN CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2000-602362/57.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; AAC75154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO200058473-A2
                                                                                                                                                                                                                                                                                                 08-FEB-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               31-MAR-1999;
02-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           30-MAR-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 05-0CT-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      05-APR-1999
                                                                                                                                                                                                                                                           AAB40945;
                                                                                                                                                                             RESULT 11
                                                                                                                                                                                              AAB40945
                                                                           ò
                                                                                                                                                                                                                                        ò
```

(first entry)

16-MAY-2001

```
26;
                                                 graft vs host disease, cardiovascular disease, diabetes mellitus, hypothyroidism, cholesterol ester storage, systemic lupus erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to enhance coagulation; to inhibit thrombosis; and as a contraceptive.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 495
                 vectors. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disorders, osteoarthritis,
                                                                                                                                                                                                                                                                                                                                                                   spsgetatsslcsvtntsmmtsekitvttstgstlgnpgetssvpvtgslmpvtsaalvt 304
                                                                                                                                                                                                                                                                                                                                                                                                        -----VGVADPNNPVFWSRDGGATWQA 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         869
                                                                                                                                                                                                                                                                                                                                 145 APSGKGLWRSTDSGATWSQMTN-----FPDVGTYIANPTDT-----TGYQSDIQGVVWVA 194
                                                                                                                                                                                                                                                                                                                                                                                                                                        305 vdpeggspatfsrtstgdttafsknhgtgsvettrvsgintlntltpv-----ttst
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VPGAPTGFIPHKGVFD---PVNHVLYIATSNTGGPY-------DGSSGDVWKFS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               vlsspsgfnpsgtvsgetfpsgettisspssvsntflvtskvfrmpisrdstlgnteets
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---TIDROHPNT--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----IMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   spgvsqeiftlhetttwpssf---sskghttwsq----telpststgaatrlvtgnpstg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GVQPNPPVPS-----LGWMDEAMAIDPFN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     aagtiprvpskvsaigepgepttysshsttlpkttgagagtgwtgetgttgeallsspsy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GGFTHADVT------AVPSTIFTSPVFTTGTSVDXAELNPSIIV------RAGSFD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        702 paptgdghttgapttalqatpsshdatlgpsggtsl---sktgaltlansvvstpggpeg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            540 VWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qwtsasastspdtaaamththqa-esteasgqtqtsepassgsrttsagtatpsssg---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     600 FQPVAAGLPSSGAVGV-----MFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----asgttpsgsegistsgettrfssnpsrdshttgsttellsasashgai-pvstgma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----SQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            svtqmikt-atspssspmldrhtsqqittap---stnhstihststspqespavsq----
                                                                                                                                                                                                                                                                                           Mismatches 280; Indels 194;
                                                                                                                                                                                                                                                        Length 1532;
 express ORFX proteins
                                                                                                                                                                                                                                                        DB 21;
                                                                                                                                                                                                                                                                             Pred. No. 0.21;
                                                                                                                                                                                                                                                          3.7%; Score 148;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VT - SGTWTRISPVPST - - - DTANDYFGYSGL - - -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SAVNVGFGKSAPGSSYPAVFVVGTIGGVTG 682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  870 ssi----vpgtfhptlseastagrptg 892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB72314 standard; Protein; 1749
                                                                                                                                                                                                                                                                                             83;
 acids can be used to
                                                                                                                                                                                                                                                                           19.38;
                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                        FDKSSSSLGQASKTIF---
                                                                                                                                                                                                     1532 AA;
                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                           Best Local Sim
Matches 133;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB72314;
                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               366
                                                                                                                                                                                                                                                                                                                                                                  245
                                                                                                                                                                                                                                                                                                                                                                                                      195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      496
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB72314
                                                                                                                                                                                                                                                                                                                                                                                                                                        සු
88888888888888888
                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

```
This invention relates to defence-related signalling genes isolated from the sunflower (Helianthus annuus). The genes encode a neoxanthin cleavage charge (NCE), an amino acid permease (AAP) and a glutamic acid rich protein (NCE). The signalling gene is useful for increasing the resistance of a plant to a pathogen such as fungus, virus, bacterium, nematode or insect (e.g. European corn borer), preferably sclerotinia spp., Phoma spp., or Phomopsis spp. by stably incorporating a construct containing the gene into the genome of the plant. The gene is useful for regulating gene expression in a plant, in response to a stimulus such as infection with a pathogen, damage from a pathogen, hydrogen peroxide, jasmonic acid, methyl jasmonate, salicylic acid, oxalic acid or expression of a gene encoding oxalic acid oxidase. The genes are also useful for stem preferred regulation of gene expression in a plant. The genes are useful in proved agronomic traits, for modifying a plant. The genes are useful in proved agronomic traits, for modifying abscisic acid (ABA) metabolism and for modifying amino acid transport and content in plants. The present sequence represents a glutamic acid rich protein (GRP) like protein isolated from Synechocystis sp. The protein is such abe.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               35;
                                                                                                                                         enzyme;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Novel isolated defence-related signalling gene isolated from sunflower encoding neoxanthin cleavage enzyme, amino acid permease or glutamic acid-rich protein useful for increasing resistance of plant to a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      172 TYIANPIDITGYQSDIQGVVWVAFDKSSSSL----GQASKT-----IFVGVADPNNPVFW 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----GSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVA-TQISW 321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         119 GGN-----MPGRGM--GERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVG 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   219 apggnitvaavpgqnririsqagsllslevevspqmnqggsfsvldlptlltqgasnldl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SRDGGATWQAVPGAPTGFIPHKG------VFDPVNH-----VLYIATSNTGGPYD---
                                                                                                                              Defence-related signalling gene; sunflower; neoxanthin cleavage NCE; amino acid permease; AAP; glutamic acid rich protein; GRP; pathogen resistance; abscisic acid metabolism.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   188;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3.6%; Score 146; DB 22; Length 1749; 21.8%; Pred. No. 0.34; Live 81; Mismatches 306; Indels 188
                                                           Glutamic acid rich protein-like protein amino acid sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Example 1; Page 128-135; 135pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ľ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (PION-) PIONEER HI-BRED INT INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hu X,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             18-AUG-1999; 99US-0149656.
23-MAY-2000; 2000US-0206405.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          17-AUG-2000; 2000WO-US22961
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Crasta OR,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (CURA-) CURAGEN CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2001-211215/21.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1749 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                              Synechocystis sp
                                                                                                                                                                                                                                                                                                                                                 WO200112801-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                       22-FEB-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Simmes 160;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DĽ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 pathogen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bidney
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   223
80 X C C C C C C C C C C C C C C C X 8 X X B X X B X X B X X B X X B X X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
```

```
New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell
                                                                                                          526
                                                                                                                                                                                                                               535
                                       381
                                                                                                                                                431 -AVNDLISPPSGAPLISALGDLGGF------THADVTAVPSTIFTSPVFTTGTSVD 479
                                                                                                                                                                                                                                                          GGTVAAS-----ADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQI-- 572
                                                                                                                                                                                                                                                                                                                                                                                                                         SYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTN- 725
                                                                                                                                                                                                                                                                                                                                                                                                                                                    707 istgsivtspfggnagnvilnaggt--lttgyietsgtnggdvttssgsntstayidtrg 764
589 fddgagittplksfgvtateiyvgndivtggnqifdgvfyglgpvnltssagsviftnni
                                                                                                                                                                                                                                                                                                                                                                      -----SAPGS
                                                                                                                                                                                                                                                                                                                                                                                      ----YAELNPSIIVRAG-SFDPSSQPN-----DRHVAFSTDGGKNWFQGSEPGGVTT
                                                                                                                                                                                                                               ngnvilgadgsiavganfsadpsapgnltfdsptitidalfslgtgsiifantgp--int
                                                                                                                                                                                                                                                                                                                573 ---RSDRVNPKTFYALSNGTFYRSTD----GGVTFQPVAAGLP----SSGAVGVMFHAVP
                                       WPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWM
                                                                                                                                                                     gakslmvddtarvntfatmgelgtwlldpleiivgttddlladpklv---svltittsld
                                                                  ---vriggdfqgq--lt-----lpnasqtli
                                                                                                                                                                                                                                                                          DEAMAIDPFNSDRMLYGTGATLYA-TNDLTKWDS------GGQIHIAPMVKGLEET-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Drosophila; developmental biology; cell signalling; insecticide;
                                                                                                                                                                                                                                                                                                                                                                      GKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAV----NVGFGK--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Drosophila melanogaster polypeptide SEQ ID NO 5061
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Myers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ABB59423 standard; Protein; 1472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PWD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    23-MAR-2000; 2000US-191637P.
11-JUL-2000; 2000US-0614150.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23-MAR-2001; 2001WO-US09231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                 qgatvdvsgnggggt----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ξ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | |: :||| |
765 fgdgleidslggavs 779
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -GRGIVYGDIGGAPS 739
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Drosophila melanogaster
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Adams M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; 2001-656860/75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (PEKE ) PE CORP NY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-PSDB; ABL03526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200171042-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  pharmaceutical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JĊ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 27-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABB59423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Venter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                726
                                                                                                                                                                           421
                                                                                                                                                                                                                                                         527
                                                                                                                                                                                                                                                                                                                                                                     622
                                                                                                                                                                        g
                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                  ď
                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                               ò
                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                   ò
                                       ò
                                                                                          ò
```

```
45;
                                                                                          įs
                                                                The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     353 tavpaaagteisa-ppglehqlvqqgshleessssgpaavtp-patlsgsattpllqysa 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                411 avsnpppglgsggtgsgagtgasaaagggagstpssfvsaspdtfssaasaaatlvhgag 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               237 ggggrgggfvsrsgrgggrmggrtggprgdrgsggpggay----gsgrggnanedhheve 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             KGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELN 484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             471 kqqqlqqqttpikpsatlsveqsqyfnslasq-----gvspgsvpvqsapagyaq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   714 sgaskta-----sgfpttap----ggyssstyanvgssvans------ygqqg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 79 NKVWAAVGMYTNSWDP-----NDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IQGVVWVAFD--KSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVP--GAPTGFIP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----SGLTIDROHPNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    313 IMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----dygsksgg-----gggvtlaaglg----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       --ssqisdalnaagytsrstsqqqqqyssavnatidqltksdpygqtggsgnayqnayqs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PSIIVRAGSFDPSSQPN-----DRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FVWAPGDP-----GQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       --IAADPINT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AVDPNNDNILYFGAPSGKGLWRSTDS-GATWSQMTNF----PDVGTYIANPTDTTGYQSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----RI---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----npv--aaysqtstsvgvsqypntyanvfasgtaagagtaegsqqqqqirrarvkl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----SGGQIHIAPMV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----ygsyqpssyqqagsgaqsgtgavsgggggtatqnipvggsssqnstsgnassay
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ST---DGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAIT
                                                                                                                                                                                                                                                                                                                                                               Length 1472;
                                    Disclosure; SEQ ID NO 5061; 21pp + Sequence Listing; English
                                                                                                                                                                                                                                                                                                                                                                                                                                             14 GGGGFVDGIVFNEG-APGILYVRT-----DIGGMYRWDAANGRW---
                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HKGVFDPVNHVLYIATSNTGG--PYDGSSGDVWKFSVTSGTWT----
                                                                                                                                                                                                                                                                                                                                                                 DB 22;
                                                                                                                                                                                                                                                                                                                                                               3.6%; Score 144.5; DB 22; 21.2%; Pred. No. 0.34; Live 91; Mismatches 342;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   373 VPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATND-LTKWD-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----WVGWNNWGYNGVVS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ---SPVPST----DTANDYFGY----
                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ---pdbbbbb----
                                                                                                                                                                                                                                                                                                        1472 AA;
                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
Les 191; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       54 LLD-----
  interactions
                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                       Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ğ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

1532 AA;

```
invention.
                                   Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           695
                                                                                                                                                                                                                                                                                                                                                       193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR47578
ID AAR4
 SXC
                                                                                                                                                     à
                                                                                                                                                                                   g
                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                       D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 The invention relates to isolated polynucleotides encoding bone marrow tissue polypeptides. The sequences of the invention can be used in the tissue polypeptides. The sequences of the invention can be used in the tissue polypeptides. The sequences of the invention as parkinson's disease, Alzheimer's disease, Huntington's disorders and amyotrophic lateral sclerosis, peripheral nervous system disorders such as peripheral nervous system disorders such as peripheral characteristic, bone degenerative disorders, periodontal cisease, cerebrovascular disorders such as stroke, viral infections, bacterial infections, fungal infections, autoinfmune disorders such as rheumatoid arthritis, multiple sclerosis, asthma and ecsema, inflammatory disorders such as crohn's disease and nephritis, hyperproliferative disorders such as sporiasis, cognitive disorders, depression and cancer (e.g. tumours of the stomach, colon, panoreas, liver, bladder, prostate, ovary, kidney, brain and skin). The sequences are also useful in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration and in healing of wounds, burns, incisions and ulcers. Sequences
                                                                                                                                                                                                                                                                                                                                                                    peripheral nervous system disorder; myeloid cell disorder; osteoporosis; lymphoid cell disorder; osteoarthritis; bone degenerative disorder; skin; periodontal disease; cerebrovascular disorder; viral infection; cancer; bacterial infection; fungal infection; autoimmune disorder; pancreas; hyperproliferative disorder; cognitive disorder; depression; stomach; inflammatory disorder; toolon; liver; bladder; prostate; kidney; ovary; brain; skin; bone; cartilage; tendon; ligament; nerve;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Drmanac RT;
saahydsvstssgvssnsgstgnggv----vsgqtganqaavsnn--nsvsgsssvsnvt 916
                                                                                                                                gttaavlasltnkntsssnssgsggsaatttgnasgggagastggvgsssg-aggagsg 1032
                                                                 917 agvasgnvagvgggvsgsgvssgvgvpggsassvgvnvnnnsssassvgaatvaqta--t 974
                                                                                                  GITWVLI------NDDQHQYGNWGQ-AITGDHANLRRVYIGTNGRGIVYGDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polynucleotides encoding bone marrow tissue polypeptides, useful for preventing, diagnosing and treating osteopathic diseases -
                                                                                                                                                                                                                                                                                                                                                       bone marrow; central nervous system disorder; haematopoiesis;
                               ------GVSSAVN---VGFGKSAPGSSYPAVFV-----VGTIGGVTGAYRSDDC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        tissue growth; tissue regeneration; wound; burn; incision; ulcer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Liu C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Yang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Asundi V,
                                                                                                                                                                                                                                                                                                                        Human bone marrow tissue polypeptide #34.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 10; Page 117-118; 202pp; English.
                                                                                                                                                                                                                 AAU71956 standard; Protein; 1532 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Tang YT,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18-APR-2000; 2000US-0522929.
24-OCT-2000; 2000US-0695783.
30-NOV-2000; 2000US-250583P.
26-JAN-2001; 2001US-0770160.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18-APR-2001; 2001WO-US12607
                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2001-626527/72.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ford JE, Boyle BJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (HYSE-) HYSEQ INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAS96248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO200179447-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Homo sapiens
                                                                                                                                                                                                                                                                                      26-FEB-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25-OCT-2001.
                                                                                                                                                                                                                                                     AAU71956;
863
                                                                                                                                                                                                                                   셤
                                                                                                                                셤
                               ò
                                                                                               õ
```

```
47;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       471 gtvesdgstestgrleekgtgesgsrdrrkidghtllgsivnrtdldprvlsnsgwgqtp 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           742 kpvsspdwnkggdivgswgippatgkppgtgw-lggplp-apakeeeptgweepspesir 799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----IIND-----DQH--QYGNWGQA-----ITGDHANLRRVYIGTNGRG 728
                                                                                                                                                            64 GYNGVVSIAADPINTNK --- VW--AAVGMYINSWDPNDGAILRSSDQGATWQITPLPFKL 118
                                                                                                                                                                                                                                             357 gvngpmgtnfg-vntnkgggvwesgaansgstswgsgngansggsrrg--wg-tpaq-nt 411
                                                                                                                                                                                                                                                                                                                                                                                                          412 gtnlpsvewnklpsnghsndsangngktftng-wksteeedggsatsgtnegssvwaktg 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 266 DG-SSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPD 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          385 MAID--PFNSDRMLYGTGATLYATNDL--TKWDSGGQIHIAPMVK----GLEETAVNDLI 436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           437 SPPSGAPLISALGDLGGFTHADVTA----VPSTIFTSPVFTTGTSVDYAELNPSIIVRAG 492
                                                                                                                                                                                                                                                                                                                         119 GGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRST---DSGATWSQMTNFPDV----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----VAFDKSSSSLGQ-----ASKTIFVGV------ADPN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           693 1-----gktssftwgnninpnnssgw--desskptpsgg-wgdppksngslgwgdss
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --tpasaisnkeassgsgwgepwgepstpattvdngtsa-----w----gkp---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          888 -idsgpsw----gepiaa----asststwgsssvgpgalsksgpksmqdgwcgddmplp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           993 inpfykgfsnisfsrdspeenvgsnkmdlsgeypsi-----eeapevssawve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       531 ikgntawdtetsprgerktdngteawgssatgtfnsgacidktspngndtssvsgwgdpk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    218 NPVFWSRDGGATWQA---VPGAPTGFI-----PHKGVFDPVNHVLYIATSNTGGPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 644 dgqkssqgwsvs-asdnwgetsrnnhwgeankksssggsdsdr----svsgwne
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  325 TIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            493 SFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         553 AVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVAAG----LP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           938 gnrptg--weeeedveigmwnsnss---gelnsslnwppytkkmsskgmmkggnkqeeaw
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           --GSSYPAVFVVGTIGGVTGAYRSDDCGTTWV-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        800 rkmeiddgtsawgdpskynyknvnmwnknvpn-----gnsrsdggagvhgll----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          609 SSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITG--------
                                                                                Indels 271;
    Length 1532;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          171 GTYIAN-PTDTTGYQSD--------IQGVV----
    DB 22;
Query Match 3.6%; Score 143.5; DB 22; Best Local Similarity 19.8%; Pred. No. 0.42; Matches 167; Conservative 94; Mismatches 312;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR47578 standard; Protein; 894 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           651 ----VSSAVNVGFGKSAP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1101 smfg 1104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     729 IVYG 732
```

```
39;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The flocculation protein is used in a method to immobilise enzymes to a microbial cell wall. The coding sequence (FLO 1) is used in the production of a recombinant polynucleotide which comprises a structural gene encoding a protein with catalytic activity and at least part of a gene encoding at least the C-terminus of a protein capable of anchoring in a eukaryotic or prokaryotic cell wall. The anchoring fragment or protein is selected from alpha agglutinin, AGA 1, FLO 1, major cell wall protein of lower eukaryotes or a proteinase of lactic acid bacteria. The recombinant polynucleotide preferably also comprises a sequence encoding a signal peptide to ensure secretion of the expressed product. The signal peptide to protein, alpha factor, alpha-agglutinin, invertuse or inulnase, alpha-amylase of Bacillus or proteinases of lactic acid bacteria. The host microorganism can be used for performing enzymatic processes on an industrial scale.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   53 tysnaaymaygya----sktklgsvggqtdisidyn-----ipcvsssgtfpcpqe 99
                                                                                                                     Immobilisation; enzyme; cell wall; alpha agglutinin; ÅGA 1; FLO 1; Major cell wall protein; glycosyl-phosphatidyl-inositol; anchoring protein; alpha factor; alpha-agglutinin; invertase; inulinase; alpha-amylase; Saccharomyces cerevisiae; floculation protein; enzymatic process; fermentation; biodegradation; catalysis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Immobilisation of enzymes to microbial cell wall - by prodn. of fusion protein of enzyme linked to anchoring protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels 224;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 3.5%; Score 143; DB 15; Length 894; Best Local Similarity 20.5%; Pred. No. 0.22; Matches 161; Conservative 94; Mismatches 308; Indels 22
                                                                                        Flocculation protein of Saccharomyces cerevisiae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Verrips CT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Example 10; Page 59-64; 99pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Toschka H,
                                                                                                                                                                                                                                                                                                                                                                 93WO-EP01763
                                                                                                                                                                                                                                                                                                                                                                                                   92EP-0202080
92EP-0203899
                                                    (first entry)
                                                                                                                                                                                                                                                       Saccharomyces cerevisiae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Schreuder MP,
                                                                                                                                                                                                                                                                                                                                                                                                                                                      (UNIL ) UNILEVER NV. (UNIL ) UNILEVER PLC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1994-035071/04.
N-PSDB; AAQ54029.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      894 AA;
                                                                                                                                                                                                                                                                                                                                                               07-JUL-1993;
                                                                                                                                                                                                                                                                                                                                                                                                   08-JUL-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                      14-DEC-1992;
                                                    19-JUL-1994
                                                                                                                                                                                                                                                                                         WO9401567-A.
                                                                                                                                                                                                                                                                                                                           20-JAN-1994
               AAR47578;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Klis FM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
```

306 355 409 522 --DISAEPWL-TFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKW 412 DSGGQIHIAPMVKGLEETAVNDLISPPSGA------PLISALGDLGGFTHADVTAV 462 633 654 634 -GLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVF--VVGTIGGVTGAYRSDDCG 690 704 SSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLXIA 257 298 tgtngvptdetvivirtptsegli---stttepwtgtftststevtt-itgtnggptdet 353 553 svssssfissessksp---tysssslplvtsattsgetassl-ppatttktsegttlvt- 607 583 YALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPG--KEGDLWLAASS----------yavsttttttepwtgtftststemttv YFGA----PSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDK 258 TSNTGGPYD-----GSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTID------RQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVL----354 vivirtptseglistttepwtgtftstste----mttvtgtnggptdetvivirtptseg PSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPG **CVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTF** -----vtsceshvctesispaivstatv----tvsgvtteyttwcpisttettkgt -----ivvktpttaissslsssssgqitssitssrpiitpfypsng-tsvisssv --ddlsgsnct----vpdpsn---691 TTWVLIN 697 ttwcpis 711 307 142 256 356 413 445 463 523 705 198 g рp g g δ δ qq à à qq QΥ ò g ò 셤 ö g ŏ g ò

Search completed: July 2, 2002, 09:15:50 Job time: 238 sec

94 156

a

õ

δ

60 --WNNWGYNGVVSIAADPINTNKV--WAA--VGMYTNSWD------

us-09-917-376-3.rai

```
GENERAL INFORMATION:
APPLICANT: MYERS, ALAN M.
APPLICANT: JAMES, MARTHA G.
TITLE OF INVENTION: ISOLATION OF SU1, A STARCH DEBRANCHING
TITLE OF INVENTION: ENZYME, THE PRODUCT OF THE MAIZE GENE
TITLE OF INVENTION: SUGARY 1
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       E: Weingarten, Schurgin, Gagnebin and Hayes LLP
Ten Post Office Square
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTESO Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/410,784A
FILING DATE: 24-MAR-1995
CLASSIFICATION: 800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       34,346
3R: ISU-002XX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME: Helie, Ph.D., Holliday C-REGISTRATION NUMBER: 34,346
REGISTRATION NUMBER: 1SU-002;
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-2290
TELEFAX: 617-451-0313
                                                                                                                                                                                                                                                                                                                                               Sequence 5, Application US/08410784A Patent No. 5912413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH: 772 amino acids
TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-terminal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskett
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MOLECULE TYPE: peptide HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FRAGMENT TYPE: 1
ORIGINAL SOURCE:
US-08-410-784A-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Boston
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STATE: MA
                                                                                                                                                                                                                                                                                                                                  US-08-410-784A-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STREET:
Sequence 29, Appl
Patent No. 5268290
Sequence 8, Appli
Sequence 15, Appl
Sequence 6, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 6, Appli
                                                                                       2, 2002, 09:12:46; Search time 39.87 Seconds (Without alignments) 453.347 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Patent No. 5457037
Patent No. 5457037
Sequence 7, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 5, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 8, Al
Sequence 44, A
Sequence 4, Al
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 22, Sequence 15, Sequence 25, Sequence 25, Sequence 25, Sequence 25, Sequence 29, Sequence 29, Sequence 29,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 11,
Sequence 5, Au
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 7
                                                                                                                                                                            1 ATTQPYTWSNVAIGGGGFVD......rIGTNGRGIVYGDIGGAPSG
                                                                                                                                                                                                                                                                                                                                                                                                                   /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PeTUS_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PeTUS_COMB.pep:*
           GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         S-09-147-236-5
S-08-194-290-7
S-08-928-361B-11
S-08-928-361B-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-346-237-8
US-09-136-574A-44
US-09-346-237-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-352-159-25
US-09-352-168-25
US-08-570-311-14
US-09-352-159-29
US-09-352-168-29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-600-452A-8
US-09-540-245A-15
                                                                                                                                                                                                                                                                            hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-362-525-22
US-08-971-692-15
US-09-296-284-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -08-322-760A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -08-325-267A-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -08-386-727-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-346-237-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-296-284-4
                                                                                                                                                                                                                                                 231628 seqs, 24425594 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUMMARIES
                                                                                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                   OM protein - protein search, using sw model
                                                                                                                                                                                                         BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5268290
                                                                                                                                                                                                                                                                                                                                                                                                       Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                        Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                US-09-917-376-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        894
720
10000
10000
12000
12000
866
866
866
11000
1000
1000
11000
11000
11000
11000
11000
11000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query
Match I
                                                                                             July
                                                                                                                                                                                                                                                                              οţ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    151.5
148.5
148.5
143.5
143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      141
140.5
140.5
136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      130
128
128
                                                                                                                                                                  Perfect score:
                                                                                                                                                                                                         Scoring table:
                                                                                                                                                                                                                                                                            Total number
                                                                                                                                                                               Sequence:
                                                                                                                                                                                                                                                 Searched:
                                                                                                                                                                                                                                                                                                                                                                                                      Database
                                                                                             Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Result
No.
```

```
Appli
, Appl
                                                                                                                                                                                                                                                                                             Appl
Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Appli
Appli
Appli
Appli
Appli
Appli
                                                                                                                                                                                                           Sednence Sed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sequence
US-09-040-005-2
US-09-52-217-115
US-09-142-623-11
US-08-131-185-48
US-08-459-499-9
US-08-459-499-12
US-08-641-623-16
US-09-082-16
US-09-661-052-16
US-08-661-052-16
US-08-476-519-11
PCT-US95-09323-11
US-08-476-519-2
PCT-US95-09323-2
US-07-689-008-2
US-09-147-236-4
```

## ALIGNMENTS

```
76 INTNKVWAA-----
                                                                                                                                                                                                                                                                                        NAME/KEY: PEPTIDE LOCATION: (1)..(7
                                                                                                                                                                                                            TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                            42;
                                                                                                                                                                                                                                                                         165 TNFPDVGTYIANPTDTTG------YQSDIQGVVWVAFDKSSSSLG----- 203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----DTANDYF-GYSGLTIDRQHPNTI---MVATQISWWPDTIIFRSTD 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          437 NRILREFTVRPAAGGTVWICLRNLGPSAATR--TSWVDSRRVVRVEWSVPRQLRQAQNEL 494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    436 ISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVR---- 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DP----GQPVVYAVGFGN-----SWAASQG---VPANAQIRSDR---VNPK 580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                641 GGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVV--GTIGGVTGAYRSDDCGTTWVLIND 698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           684 NYAIAYAINGPSL------GDSNDSIYVAXNGWSSSVTFTLPAPPSGTQWYRVTD 732
                                                                                                                                                                                                                                    QGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQM 164
                                                               Gaps
                                                                                                                     WIPLLDWVGWNNWGYN---GVVS---IAADPINTNKVWAAVGMYTNSWDPNDGAILRSSD 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               339 IYSWRGLDNATYYELTSGNQYFYDNTGIGANFNTYNTVAQNLIVDSVAYWANTM---GVD
                                                                                                                                                                                                                                                                                                                                                                                                          --YAPKGVVLVPSTQSTGTKPTRAQKDDVIYEVHVRG----FTEQDTSIPAQYRGTYYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---QASKTIFVGVA------DPNNPV------FWSRDGGATWQA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | : | | : | : | | : | : | | : | : | | : | : | | : | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   233 VPGAPTG-FIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        396 GFR-----FDLASVLGNSCLNAVHASA------PNCPNGGYNFDAADSNVAI----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 393 DRMLY-----GTGATLY-----ATNDLTKW-DSGGQIHIAPMVKGLEETAVNDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     --AGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTV---AASADGSRF-VWAPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    597 TPLMQGGDEYLRTLQCNNNAYNLDSSANWLTYSWTTDQSNFYTFAQRLIRSARHIPLRPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            581 TFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTN
                                                            Indels 293;
                                                                                                                                                                        108 WGP--NWPYASNWGKGSQAGCVSDVDANGDRFNPNKLLLDPYAQEBSQDP-----
   Length 772;
   2;
; Score 162.5; DB 2
; Pred. No. 0.00071;
90; Mismatches 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              699 DQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDIG 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        T----CDWNDGASTFVAPGSETLIG---GAGTTYGQCG 763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: Bisgard-Frantzen, Henrik
APPLICANT: Svendsen, Allan
TITLE OF INVENTION: Starch Debranching Enzymes
FILE REFERENCE: 5629.200-US
CURRENT APPLICATION NUMBER: US/09/346,237A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 8, Application US/09346237A
Patent No. 6265197
GENERAL INFORMATION:
   1.08;
Similarity 19.88;
                                                            Conservative
                                                            162;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-346-237-8
Query Match
Best Local S.
Matches 162
                                                                                                                                                                                                                                 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               545
                                                                                                                                                                                                                                                                                           156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                491
                                                                                                                                                                        g
                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              α
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

```
445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        660 GNVMEOLRWFKPDGGVADATYFNDANNHAIAWRIDGSEFGDTASAIYVAHN---AWSA-- 714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     340 DKTTYNITSFRGLDNPTYYSLTADFONSWDNTGVGGNYNTRNTIAONLIVDSLAYWRDKL 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        440 --SGAPLIS---ALG----DLGGF-----THADVTAVPSTIFTSPVFTTGTSVDYA 481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---VGMYTNSWDPNDGAILRSSDQGATWQITPL 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  115 PFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYI 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  220 VFWSRDGGATWQAV---PGAPTGFI------PHKGVFDPV---NHVLYIATSNTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  600 LRSMNCNNNPYNLDSSANWLNWSRTTDQNNFQSFSKAMIAFRKAHPALRPANFYSSVDNN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            94 P---TSTLSGQGLG-----GTLYYGYRAWGPNW---PYNASWTKGSSL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         175 ANPTDITGYQSDIQG------VWWVAFDKSSSSLGQASKIIFVGVADPNNP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     263 -----GPYDGSSGDVWKFSVTSGTWTRISPV------PSTDTANDYFGYSGLTI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               342 DWTSYPNRSLR-----PKLCISAE-PWLTFGVQPNPPVPS-----PKLGWMDEAM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               400 GVDGYRFDLASVLGNSCQHGCFNFDKMDAG------NALNRIVAELPPRPAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     482 ELNPSIIVR-AGSFDPSSQPNDRHVAFSTDGGKNWFQG---SEPGGVTTGGTVAASADGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        538 RFVW--APGDPGQPVVYAVGFGNSW----AASQ-------GVP----ANAQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               546 NOVWPYGPSDGGE----DNNNSWDQGGIAADQRKAARNGMALMMLSAGVPMIVGGDEA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IRS-----ALSNGTFYRSTD-G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GVTFQPVAAGLPSSGAVGVMF-----HAVPGK-EGDLWLAASSGLYHSTNGGSSWSAIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  288 FAPDRRYAYDKTPGGPTREFKEMVKAFHDNGIKVLVD------VVYNHTGEGGAWSPT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     386 AIDPFNSDRMLYGTGATLYATNDLTKWDSGGIHIAPMVKGLEETAVNDLIS--PP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     446 GGSGVDLIAEPWAIGGNSYQVGGFPSGWAEWNGAYRDVVRQAQNKLGSVAITTG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels 313;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 774;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 3.9%; Score 158.5; DB 4; Best Local Similarity 18.8%; Pred. No. 0.0014; Matches 158; Conservative 103; Mismatches 265;
CURRENT FILING DATE: 1999-07-01
EARLIER APPLICATION NUMBER: PA 1998 00868
EARLIER FILING DATE: 1998-07-02
EARLIER FILING DATE: 1998-07-28
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 8
                                                                                                                                                                                                                                                                                                            ORGANISM: Favobacterium odoratum
                                                                                                                                                                                                                                                                                                                                                                                                   ; LOCATION: (1)..(774); OTHER INFORMATION: ISOamylase US-09-346-237-8
```

44;

```
1037 QIKVWYANGNLGSMTNVLNPKIKIENVGTTAVDLSRVKVRYWYTIDGEATQSVSVTSSIN 1096
                                                                                                                                                                                                                          300
                                                                                                                                                                                                                                                                                                                                           443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-----LVWGLEPGAATSTPAPTSTPTPTPTPTVTATPTPTPTPTGSPGTGSGVK 680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       917 VDLSRVKVRYWYTIDGEATQSVSVASSINPAYIDVKFVKLGANAGGADYYVEIGFKSGAG 976
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     615 VMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSA-----ITGVSSAVNVGFGKSAPGS 666
                                                                                                                                                                                                                                                                   534
                                                                                                                                                                                                                                                                                                               SGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVL-DISA 359
                                                                                                                                                                                                                                                                                                                                                                                                     360 EPWLIFGVQPNPPVPSPKLGW----MDEAMAIDPFNSDRMLYGTGATLYATNDLIKWDSG 415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -P 463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           681 VLYKNNETSASTGSIRPWFKIVNGGSSSVDLSRVKIRYWYTVDGDKPQSAVCDWAQIGAS 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   741 NVTFNFVKLSSGVSGADY-YLEVGFSSGAGQLQPGKDTGDIQVRFNKNDWSNYNQADDWS 799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         800 WLQSMTNYGENAKVTLYVDGVLVWGQEPGGATPAPTSTATPTPTATPTPTPTPTPT 859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               860 VSATPTPAPTASPVG-GSYWTPSESYGALKVWYANGNLSSPTNVLNPK--IKIENVGTTA 916
344 AAFLACVYADWSGCDSNKKTKYLNFAKSQIDYALGST--GRSFVVGFGTNYPQHPHHRN- 400
                                                                                        ----AHSSWANSMKIPEYHRHILYGALVGGPGSDDSYNDDITDYVQNE 444
                                                                                                                                                                           SYPAVEVVGTIGGVTGA----YRSDDCGTT--------WVLINDDQHQYGNWGQAIT
                                                                                                                                                                                                                                                          482 EIEVESKFGNSQGP---NYTEVISYIYNRTGW----PPRVTDKLSFKYFIDLTELIQAGY
                                                                                                                                                                                                                                                                                                                                                                                                                                               573 TK----IYPGGEVEHKKQAQFKISVPQGYPWDPTN-DPSYKGLTSQLEKNKYIAAYDNN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----GQPVVYAVGFGNSWAASQGVPA-----NAQIRSDRVNPKTFYALSN-GT--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----FYRSTDGGVT------FQP-----VAAGLPSSGA-----VG
                                              141 LYFGAPSGKGLWRSTDSGATWSQMTNFPD-----VGTYIANPTDTTGYQSDIQGVVW--
                                                                                                                                    193 VAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNH
                                                                                                                                                                                                                          VLYIAT -- SNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTAN-DYF-----GY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           416 GQIHIAPMVKGLEETAVNDLISP------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STIFTSPVFTTGTS-VDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKN----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----GSEPGGVTTGGTVAASADGSRFVWAPGDP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: Svendsen, Allan
TITLE OF INVENTION: Starch Debranching Enzymes
FILE REFERENCE: 5629.200-US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----LISALGDL------GGFTHADVTAV--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CURRENT APPLICATION NUMBER: US/09/346,237A
CURRENT FILING DATE: 1999-07-01
EARLIER APPLICATION NUMBER: PA 1998 00868
EARLIER FILING DATE: 1998-07-02
EARLIER APPLICATION NUMBER: 60/094,353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1097 PAYIDVKFVKLGANAGGADYYVEIG 1121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   712 GDHANLRRVYIGTNGRGI-VYGDIG 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 4, Application US/09346237A Patent No. 6265197 GENERAL INFORMATION: APPLICANT: Bisgard-Frantzen, Henrik
                                                                                                                                                                                445 VACDYNAGIVGALAK------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WFO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 4
US-09-346-237-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   444
                                                                                                                                                                                                                                                                                                                  301
                                                                                                                                                                                                                                                                                                                                                            535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             590
                                                                                        401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                            ò
                                                                                   g
                                                                                                                                    δŏ
                                                                                                                                                                              g
                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                               엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                            Treating Cellulose Containing Fabrics Using Truncated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            49;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                46 AANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVGM-YTNSWDPNDGAILRSSD 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      290 VRYGILIMLAKITGKEL--YKGAVERNLDHWIDRITYTPKGMAYLTGW----GSLRYATI 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -ITPLPFKLGGNMPGR----GMGERLAVDPNNDNI 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --IGGMYRW----D 45
                                                          -----QVNETLPWPGAGKSWYRVTDTCGWAEGASQVQAPGSE-ALVGGENTAYGLCGR 766
                    650 GVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQ 708
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels 351;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 1751;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Center, P.O. Box
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Cellulase Enzyme Compositions
                                                                                                                                                                                                                                                                                                                                                                                                        and Methods for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          J.012;
349;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    6 YTWSNVAIGGGGFVDGIVFNEGAPGILYVRTD-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE/DOCKET NUMBER: 1997US001/CIP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 3.8%; Score 152.5; D
Best Local Similarity 18.7%; Pred. No. 0.012
Matches 184; Conservative 101; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,574A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: US 08/932,571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-09-136-574A-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE: September 19, 1997
                                                                                                                                                                                                                                                                                                                                                                                Williams, Diane P. INVENTION: Compositions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REGISTRATION NUMBER: 31,215
                                                                                                                                                                                                                                            APPLICANT: Farrington, Graham K.
                                                                                                                                                                           Sequence 44, Application US/09136574A Patent No. 6294366 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE: 19-Aug-1998
CLASSIFICATION: <Unknown>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER: IBM Compatible OPERATING SYSTEM: DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGTH: 1751 amino acids
                                                                                                                                                                                                                                                                      Anderson, Paige
Gibbs, Moreland
Bergquist, Peter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEFAX: 215-540-5818
                                                                                                                                                                                                                                                                                                                                       Daniels, Roy
Morgan, Hugh W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INFORMATION FOR SEQ ID NO: 44: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CITY: Spring House
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mary E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            <Unknown>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NUMBER OF SEQUENCES;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---ATWQ---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STATE: PA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ZIP: 1947
                                                                                                                                                                                                                                                                                                                                                                                                     Ģ.
                                                                                                                                                  US-09-136-574A-44
                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;
90
                                                                715
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  105
                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                         ò
```

us-09-917-376-3.rai

```
엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ģ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                           42;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       183 KGVVLVPSTQSTGTKPTRAQKDDVIYEVHVRGFTEQDTSIPAQYRGTYGAGLKASYLAS 242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     349 ATYYELTSGNQYFYDNTGIGANFNTYNTVAQNLIVDSLAYWANTM---GVDGFR-----F 400
                                                                                                                                                                                                                                                                                                                                                                                                  MYTNSWDFNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDFNNDNILYFGAP 146
                                                                                                                                                                                                                                                                                                                                                                                                                                    148 AQEVSQDP-----LNPSNQ------NGNVFASGASYR-----TTDSGIY--AP 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LGQASKTIFVGVAD------PN---PNVFW-----SRDGGATWQAVPGAPTG-F 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        302 QAMVQAFHNAGIKVYMDV------VYNHTAEGGTWTSSDPTTATIYSWRGLDN 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----DIANDYF-GYSGLTIDRQHPNTI---MVATQISWWPDTIIFRSTDGGATWTRIW 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    452 GGSGLDLFA----EPWAIGGNSYQLGGFPQGWSEWNGLFRDSLRQAQNEL-----GSMTI 502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          503 YVIQDANDFSGS-----SNLFQSSGRSPWNSINFIDVHDGMTLKDVYSCNGANNSQAWP 556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SCKGLWRSTDSGATWSQMTNFPDV - GTYIANPTD-TTGYQSDIQGVVWVAFDKSS--SS 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 243 LG-VTAVEFLPVQETQNDANDVVPNSDANQNYWGYMTENYFSPDRRYAYNKAAGGPTAEF 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DWTS-YPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAID----PFNSDRML 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DLASVLGNSCLNGAYTASA------PNCPNGGYNFDAADSNVAINRILREFTVRPAA 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       397 YGTGATLYATNDLTKWDSGGQIH-IAPMVKGLEE-----TAVNDLISPPSGAPLI 445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVR------AGS 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        494 FDPSSQPNDRHVAFSTDGGK----NWFQGSEPGGVTTGGTV---AASADGSRFVWAPGDP 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  600 GTPLMOG---GDEYLRTLQCNNNAYNLDSSANWLTYSWTTDQSNFYTFAQRLIAFRKAHP 656
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -GPSLGDSNSIYVAYNGWSSSVTFTLPAPPSGTQWYRVTDT----CDWNDGASTFVAPGS 752
                                                                                                                                                                                                                                                                                                                            GAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVG 86
                                                                                                                                                                                                                                                                                                                                                            96 GITGAVY-----YGYRAMGPNWPYASNWGKGSQAGFVSDVDANGDRFNPNKLLLDPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     241 IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPST-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            557 YGPS-----DGGTSTNYSWDQGMSAG---TGAAVDQRRAARTGMAFEML--SA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                547 GQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTF-YALSNGTFYRSTDGGVTFQPVAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        606 GL-PSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNG----GSSWSAIT--GVSSAVNVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           657 ALRPSS--------WYSGSQLTWYQPSGAVADSNYWNNTSNYAIAYAIN--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLR
                                                                                                                                                                                                                                                       Length 776;
                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                       DB 4;
                                                                                                                                                                                                                                                                                         329;
                                                                                                                                                                                                                                                       3.8%; Score 151.5; DB 20.6%; Pred. No. 0.0045
                                                                                                                                                                                                                                                                                           Mismatches
                                    3.0
                 NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSEQ for Windows Version
SEQ ID NO 4
                                                                                                            amyloderamosa
                                                                                                                                                                                                                                                                                       91;
EARLIER FILING DATE: 1998-07-28
                                                                                                                                                            LOCATION: (1)..(776)
COTHER INFORMATION: Isoamylase
US-09-346-237-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RVYIGTNGRGIVYGDIG 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ETLIG -- GAGTTYGQCG 767
                                                                                                                                                                                                                                                                      Similarity 20.6
4; Conservative
                                                                                         TYPE: PRT
ORGANISM: Pseudomonas
                                                                                                                                              NAME/KEY: PEPTIDE
                                                                                                                                                                                                                                                                      Best Local Sim
Matches 164;
                                                                      LENGTH: 776
                                                                                                                                                                                                                                                       Query Match
                                                                                                                           FEATURE:
                                                                                                                                                                                                                                                                                                                            27
                                                                                                                                                                                                                                                                                                                                                                                                  87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          753
                                                                                                                                                                                                                                                                                                                                                            Db
                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            · ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ç
```

```
LG-VTAVEFLPVQETQNDANDVVPNSDANQNYWGYMTENYFSPDRRYAYNKAAGGPTAEF 275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  70 GITGAVY-----YGYRAWGPNWPYASNWGKGSQAGFVSDVDANGDRFNPNKLLLDPY 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     122 AQEVSQDP-----LNPSNQ------NGNVFASGASYR-----TTDSGIY--AP 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SGKGLWRSTDSGATWSQMTNFPDV--GTYIANPTD-TTGYQSDIQGVVWVAFDKSS--SS 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QAMVQAFHNAGIKVYMDV------VYNHTAEGGTWTSSDPTTATIYSWRGLDN 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----DIANDYF-GYSGLTIDRQHPNTI----MVATQISWWPDTIIFRSTDGGATWTRIW 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TTYYELTSGNOYFYDNTGIGANFNTYNTVAQNLIVDSLAYWANTM---GVDGFR----F 374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DWTS-YPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAID----PFNSDRML 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DLASVLGNSCLNGAYTASA-----PNCPNGGYNFDAADSNVAINRILREFTVRPAA 425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                397 YGTGATLYATNDLTKWDSGGQIH-IAPMVKGLEE------TAVNDLISPPSGAPLI 445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GGSGLDLFA----EPWAIGGNSYQLGGFPQGWSEWNGLFRDSLRQAQNEL----GSMTI 476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         477 YVTQDANDFSGS-----SNLFQSSGRSPWNSINFIDVHDGMTLKDVYSCNGANNSQAWP 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FDPSSQPNDRHVAFSTDGGK----NWFQGSEPGGVTTGGTV---AASADGSRFVWAPGDP 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GTPLMOG----GDEYLRTLQCNNNAYNLDSSANWLTYSWTTDQSNFYTFAQRLIAFRKAHP 630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GL-PSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNG----GSSWSAIT--GVSSAVNVG 658
                                                         GALLI, GIULIANO: LUCCHESE, GIUSEPER, PAOLO; CAMERINI, BARBARA;
GALLI, GIULIANO: LUCCHESE, GIUSEPER; GRANDI, GUIDO; DI GENNARO,
TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLASE
ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/1,797
FILING DATE: 08-JAN-1993
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: 749,621
FILING DATE: 19-AUG-1991
STELING DATE: 224,114
FILING DATE: 25-JUL-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        27 GAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LGQASKTIFVGVAD------PN---PN-----SRDGGATWQAVPGAPTG-F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             87 MYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTF-YALSNGTFYRSTDGGVTFQPVAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KGVVLVPSTQSTGTKPTRAQKDDVIYEVHVRGFTEQDTSIPAQYRGTYYGAGLKASYLAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVR------AGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     YGPS------DGGTSTNYSWDQGMSAG---TGAAVNQRRAARTGMAFEML--SA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 750;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPST-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ouery Match 3.7%; Score 148.5; DB 6; Best Local Similarity 20.6%; Pred. No. 0.0072; Matches 164; Conservative 90; Mismatches 330;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQ ID NO:3:
LENGTH: 750
5457037-3
;Patent No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              909
```

42;

S

```
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QAMVQAFHNAGIKVYMDV--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
Matches 164; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME/KEY: PEPTIDE
                 532 YGPS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQ ID NO 7
                                                                                                                 909
                                                                                                                                                                                  629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     302
                                               ò
                                                                             g
                                                                                                                 δy
                                                                                                                                              셤
                                                                                                                                                                                ò
                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     42;
                                                        672 -GPSLGDSNSIYVAYNGWSSSVTFTLPAPPSGTQWYRVTDT----CDWNDGASTFVAPGS 726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 71 GITGAVY-----YGYRAWGPNWPYASNWGKGSQAGFVSDVDANGDRFNPNKLLLDPY 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                87 MYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAP 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AQEVSQDP-----INPSNQ------NGNVFASGASYR-----TTDSGIY--AP 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -PN---NPVFW----SRDGGATWQAVPGAPTG-F 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QAMVQAFHNAGIKVYMDV-------VYNHTAEGGTWTSSDPTTATIYSWRGLDN 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DWTS-YPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAID----PFNSDRML 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DLASVLGNSCINGAYTASA-----PNCPNGGYNFDAADSNVAINRILREFTVRPAA 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVR-----AGS 493
---WYSGSQLTWYQPSGAVADSNYWNNTSNYAIAYAIN--
                              FGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 27 GAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SGKGLWRSTDSGATWSQMTNFPDV--GTYIANPTD-TTGYQSDIQGVVWVAFDKSS--SS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LG-VTAVEFLPVQETQNDANDVVPNSDANQNYWGYMTENYFSPDRRYAYNKAAGGPTAEF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ----DTANDYF-GYSGLTIDRQHPNTI---MVATQISWWPDTIIFRSTDGGATWTRIW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TTYYELTSGNOYFYDNTGIGANFNTYNTVAQNLIVDSLAYWANTM---GVDGFR----F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GGSGLDLFA----EPWAIGGNSYQLGGFPQGWSEWNGLFRDSLRQAQNEL----GSMTI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   YVTQDANDFSGS ----SNLFQSSGRSPWNSINFIDVHDGMTLKDVYSCNGANNSQAWP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FDPSSQPNDRHVAFSTDGGK----NWFQGSEPGGVTTGGTV---AASADGSRFVWAPGDP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPST
                                                                                                                                                                                          S457037-5
;Patent No. 5457037
; Patent No. 5457037
; APPLICANT: TOGNONI, ANGELO;CARRERA, PAOLO;CAMERINI, BARBARA;
;GALLI, GIULIANO;LUCCHESE, GIUSEPPE;GRANDI, GUIDO;DI GENNARO, CARLO
TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLASE
; ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 751;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3.7%; Score 148.5; DB 6;
20.6%; Pred. No. 0.0072;
Live 90; Mismatches 330;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 YGTGATLYATNDLTKWDSGGQIH-IAPMVKGLEE---
                                                                                                                                                                                                                                                                                                          CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/1,797
FILING DATE: 08-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 749,621
FILING DATE: 19-AUG-1991
APPLICATION NUMBER: 224,114
FILING DATE: 25-JUL-1988
                                                                                              RVYIGTNGRGIVYGDIG 735
                                                                                                                              ETLIG--GAGTTYGOCG 741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 3.7%
Best Local Similarity 20.6°
Matches 164; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LGQASKTIFVGVAD-----
                                                                                                                                                                                                                                                                                               NUMBER OF SEQUENCES:
ALRPSS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH: 751
                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQ ID NO:5:
                                                                                              719
                                                                                                                              727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5457037-5
                                                          g
                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
42;
673 -GPSLGDSNSIYVAYNGWSSSVTFTLPAPPSGTQWYRVTDT----CDWNDGASTFVAPGS 727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   96 GITGAVY-----YGYRAWGPNWPYASNWGKGSQAGFVSDVDANGDRFNPNKLLLDPY 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       87 MYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAP 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -- VYNHTAEGGTWTSSDPTTATIYSWRGLDN 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GAPGILYVRIDIGGMYRWDAANGRWIPLLDWYGWNNWGYNGVVSIAADPINTNKVWAAVG 86
                                                                                                                                                                                                                                                                                          632 ALRPSS-------WYSGSQLTWYQPSGAVADSNYWNNTSNYAIAYAIN--
                                                                                                                                                                                                                                                                                                                                                              FGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----DIANDYF-GYSGLTIDRQHPNTI---MVATQISWWPDTIIFRSTDGGATWTRIW
                                                                                                  GQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTF-YALSNGTFYRSTDGGVTFQPVAA
                                                                                                                                                            575 GTPLMQG---GDEYLRTLQCNNNAYNLDSSANWLTYSWTTDQSNFYTFAQRLIAFRKAHP
                                                                                                                                                                                                                             GL-PSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNG----GSSWSAIT--GVSSAVNVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        148 AQEVSQDP----TUPSNQ-----NGNVFASGASYR----TTDSGIY--AP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SCKGLWRSTDSGATWSQMTNFPDV--CTYIANPTD-TTGYQSDIQGVVWVAFDKSS--SS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KGVVLVPSTQSTGTKPTRAQKDDVIYEVHVRGFTEQDTSIPAQYRGTYYGAGLKASYLAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LGQASKTIFVGVAD-----PN---PN---NPVFW-----SRDGGATWQAVPGAPTG-F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LG-VTAVEFLPVQETQNDANDVVPNSDANQNYWGYMTENYFSPDRRYAYNKAAGGPTAEF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPST-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 776;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 148.5; DB 4;
Pred. No. 0.0075;
0; Mismatches 330;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
APPLICANT: Bisgard-Frantzen, Henrik
APPLICANT: Svendsen, Allan
TITLE OF INVENTION: Starch Debranching Enzymes
FILE REPERENCE: 5629.200-US
CURRENT APPLICATION NUMBER: US/09/346,237A
CURRENT FILING DATE: 1999-07-01
EARLIER APPLICATION NUMBER: PA 1998 00868
EARLIER APPLICATION NUMBER: 60/094,353
EARLIER FILING DATE: 1998-07-02
EARLIER FILING DATE: 1998-07-02
SARLIER FILING DATE: 1998-07-28
NUMBER: OF SEQ ID NOS: 14
SOFTWARE: FastSEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 7, Application US/09346237A Patent No. 6265197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3.7%; Scu.
20.6%; Pred
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM: Pseudomonas species
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COP-346-237-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            719 RVYIGTNGRGIVYGDIG 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             728 ETLIG--GAGTTYGQCG 742
```

φ

```
MOLECULE TYPE: protein US-08-362-525-22
                                                                                                                                                                                                            TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                           09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          583
                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ENZYMES TO THE CELL BY PRODUCING A FUSION
                                                                                                             445
                                                                                                                                                                                                546
TTYYELTSGNQYFYDNTGIGANFNTYNTVAQNLIVDSLAYWANTM---GVDGFR----F 400
                                                                              451
                                                                                                                                                                          SALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVR------AGS 493
                                                                                                                                                                                                                                                                                                            605
                                                                                                                                                                                                                                                                                                                                                                                                659 FGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLR 718
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -GPSLGDSNSIYVAYNGWSSSVTFTLPAPPSGTQWYRVTDT----CDWNDGASTFVAPGS 752
                                                                                                                                                                                                                                                                                                       547 GQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTF-YALSNGTFYRSTDGGVTFQPVAA
                                                                                                                                                                                                                                                                                                                                      600 GTPLMQG---GDEYLRTLQCNNNAYNLDSSANWLTYSWTTDQSNFYTFAQRLIAFRKAHP
                                              DWTS-YPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAID----PFNSDRML
                                                                            DLASVLGNSCLNGAYTASA - - - - - - - PNCPNGGYNFDAADSNVAINRILREFTVRPAA
                                                                                                             -----TAVNDLISPPSGAPLI
                                                                                                                                   FDPSSQPNDRHVAFSTDGGK----NWFQGSEPGGVTTGGTV---AASADGSRFVWAPGDP
                                                                                                                                                                                                                                                                        557 YGPS------DGGTSTNYSWDOGMSAG---TGAAVDORRAARTGMAFEML--SA
                                                                                                                                                                                                                                                                                                                                                                          606 GL-PSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNG----GSSWSAIT--GVSSAVNVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATE: SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE OF INVENTION: PROCESS FOR IMMOBILIZING TITLE OF INVENTION: CELL WALL OF A MICROBIAL TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CUSHMAN DARBY & CUSHMAN, L.L.P.
                                                                                                         397 YGTGATLYATNDLTKWDSGGQIH-IAPMVKGLEE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 92202080.5
FILING DATE: 08-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 92203899.7
FILING DATE: 14-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP93/01763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CURRENT APPLICATION DATA:
CURRENT APPLICATION NUMBER: US/08/362,525
FILING DATE: 04-JAN-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1100 New York Avenue, N.W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-362-525-22
; Sequence 22, Application US/08362525
; Patent No. 6027910
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: KLIS, FRANCISCUS M.
APPLICANT: SCHREUDER, MAARTEN P.
APPLICANT: TOSCHKA, HOLSER Y.
APPLICANT: VERRIPS, CORNELIS T.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RVYIGTNGRGIVYGDIG 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       753 ETLIG -- GAGTTYGQCG 767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STATE: D. C. COUNTRY: U.S.A. ZIP: 20005-3918 COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CORRESPONDENCE ADDRESS: ADDRESSE: CUSHMAN D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE: 04
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STREET:
CITY: Wa
                                                                                                                                                                            446
               349
                                            342
                                                                            101
                                                                                                                                                                                                                                         494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      719
             ద
                                                                                                                                                                                                                                                                        엄
                                            å
                                                                         8
                                                                                                        ò
                                                                                                                                      g
                                                                                                                                                                          ò
                                                                                                                                                                                                       a
                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

```
409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100 DSYGNWGCKGMGACS----NSQGIAYWSTDLFGFYTTPTNVTLEMTGYFLPPQTGSYTFK 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        156 FATVDDSAIL--SVGGATAFNCCAQQQPPITSTNFTIDGIKPWGG-----SLPPNIEGTV 208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 634 -GLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVF--VVGTIGGVTGAYRSDDCG 690
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              410 LVTTTTEPWTGTF-----------------TSTSTEMSTVTGTNGLPTDETV-- 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           553 SVSSSSFISSESSKSP---TYSSSSLPLVTSATTSQETASSL-PPATTKTSEQTTLVT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      53 TYSNAAYMAYGYA----SKTKLGSVGGQTDISIDYN-----IPCVSSSGTFPCPQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            94 ---PNDGAILRSSDQGATW------QITPLPFKLGGNMPGRGMGERLAVDPNNDNIL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           142 YFGA----PSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             198 SSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          354 VIVIRTPISEGLISTITEPWIGTFISISTE----MITVIGINGOPIDETVIVIRTPISEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           --DISAEPWL-TFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                445 ------IVVKTPTTAISSSLSSSSGQITSSITSSRPIITPFYPSNG-TSVISSSV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       463 PSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 655 KGTTEQTTETTKQTTVVTISSCESDVCSKTAS----PAIVSTSTATINGVTTEY----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     209 YMYAGYYYPM-KVVYSNAVSWGTLPISVTLPD-GTTVSD-------DFEGYVY-SFD-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   258 TSNTGGPYD-----GSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTID--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GVTTGGTVAASADGSREVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----VTSCESHVCTESISPAIVSTATV-----TVSGVTTEYTTWCPISTTETTKQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --WINWGYNGVVSIAADPINTNKV--WAA--VGMYTNSWD------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 413 DSGGQIHIAPMVKGLEETAVNDLISPPSGA------PLISALGDLGGFTHADVTAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----RQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVL-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            YALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPG--KEGDLWLAASS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        224;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       7 TWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDMVG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 894;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3.5%; Score 143; DB 3; Length 89.
20.5%; Pred. No. 0.023;
Live 94; Mismatches 308; Indels
                                                                                                                                                           213289/T7020(V)
EILING DATE: 07-JUL-1953
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 2132F
TELECOMMUNICATION INFORMATION:
TELEPAX: (202) 821-9940
TELEXX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 894 amino acids
:: ENGTH: 894 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 3.5%;
Best Local Similarity 20.5%;
Matches 161; Conservative
```

39;

```
705 TTWCPIS 711
                                                                                                                                                                                                                                           691 TTWVLIN 697
                                                                                                                                                                                                                                                                                                                                       US-09-296-284-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-296-284-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQ ID NO 25
LENGIH: 720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                              608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      355
                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                           g
                                                                                                              δλ
                                                                                                                                             g
                                                                                                                                                                           ò
                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    39;
                                                                                                                                                                        TITLE OF INVENTION: Immobilized proteins with specific binding TITLE OF INVENTION: capacities and their use in processes and products. NUMBER OF SEQUENCES: 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        298 TGINGVPIDETVIVIRIPISEGLI---STITEPWIGTFISTSTEVIT-ITGINGQPIDET 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            354 VIVIRTPTSEGLISTTTEPWIGTFISTSTE----MITVTGINGQPTDETVIVIRTPTSEG 409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   142 YFGA----PSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDK 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIA 257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----TSTSTEMSTVTGTNGLPTDETV-- 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  413 DSGGQIHIAPMVKGLEETAVNDLISPPSGA------PLISALGDLGGFTHADVTAV 462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----IVVKTPTTAISSSLSSSSGQITSSITSSRPIITPFYPSNG-TSVISSSV 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7 TWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVG----- 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --DDLSQSNCT----VPDPSN------YAVSTTTTTTTTTTSTSTSTEMTTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---PNDGAILRSSDQGATW-----QITPLPFKLGGNMPGRGMGERLAVDPNNDNIL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          156 FATVDDSAIL -- SVGGATAFNCCAQQQPPITSTNFTIDGIKPWGG-----SLPPNIEGTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TSNTGGPYD------GSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTID--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --DISAEPWL-TFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----RQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 3; Length 894;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                       COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/971,692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --WNNWGYNGVVSIAADPINTNKV--WAA--VGMYTNSWD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 3.5%; Score 143; DB 3; L Similarity 20.5%; Pred. No. 0.023; 1; Conservative 94; Mismatches 308;
                                                                                                           Sequence 15, Application US/08971692 Patent No. 6114147 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                   CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LVTTTTEPWTGTF------
                                                                                                                                                                                                    NUMBER OF SEQUENCES: 40
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                     : 894 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                    protein
691 TTWVLIN 697
                          705 TIWCPIS 711
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; MOLECULE TYPE:
US-08-971-692-15
                                                                                                                                                                                                                                                                                                                                       FILING DATE:
                                                                             RESULT 9
US-08-971-692-15
                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches 161;
                                                                                                                                                                                                                                                                                                                                                                                                       LENGIH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    445
                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ġ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
```

```
43;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Choi, Eui-Sung
APPLICANT: Rhee, Sang-Ki
APPLICANT: Rhee, Sang-Ki
APPLICANT: Leb, Eun-Hae
TITLE OF INVENTION: Gluconobacter Suboxydans Sorbitol Dehydrogenase, Genes
TITLE OF INVENTION: and Methods of Use Thereof
FILE REFERENCE: 1533.0870000
CURRENT APPLICATION NUMBER: US/09/296,284A
CURRENT FILING DATE: 1999-04-22
SOFTWARE: Patentin Ver. 2.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      494 ISSSUTSSLETSSPUISSSUITSTITSTIFSESKSSVIPTSSSTSGSSE-SETSSAG 552
                                                                                                                                                                 607
                                                                                                                                                                                                                                                              -----VISCESHVCTESISPAIVSTATV----TVSGVTTEYTWCPISTTETTKQT 654
                                                                                                                                                                                                                                                                                                                                   634 -GLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVF--VVGTIGGVTGAYRSDDCG 690
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              76 INTNKVWAAVGMYTNSWDPNDGAI--LRSSDQGATWQITPLPFKLGGNMPGRGMGERLAV 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              134 DPNNDNILYFGAPS-GKGLWRSTDSGATWSQMTNFPDVGTYIANPTDFTGYQSDIQGVVW 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              250 VNHVLXIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQH 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    310 PNTIMVATQISWWPDTIIFRSTDGGATW-----TRIWDWTSYP------NRSLRYV 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LDISAEPWLIFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDS 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           26 EGAP----GILYVRTDIGGMYRWDAANGRWIPLLDWVGWNWGYNGVV--SIA----ADP 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            193 VAF---DKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDP
                                                                                                                                                                                                                                                                                                                                                                   463 PSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPG
                                                                                                          GVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTF
                                                                                                                                                583 YALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPG--KEGDLWLAASS----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3.5%; Score 141; DB 4; Length 720;
20.4%; Pred. No. 0.024;
ive 73; Mismatches 235; Indels 258;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 25, Application US/09296284A Patent No. 6204040 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM: Gluconobacter suboxydans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 20.49
Matches 145; Conservative
```

:: :: ::

```
35 LDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDS 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            140 ---GGALEYGWNSFHLYDGASFAANQDVIAVTINYRTNI-----LGFPAAPQLPITQRNL 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 3.5%; Score 140.5; DB 4; Best Local Similarity 20.9%; Pred. No. 0.041; Matches 159; Conservative 69; Mismatches 224;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-352-159-25
; Sequence 25, Application US/09352159A
; Patent No. 6211434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; LOCATION: (1)...(24)
US-09-352-159-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM: Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAME/KEY: SIGNAL
                                                                                                                                                                                                                                                                                                                     579
                                          361
                                                                                                                                                                                                                                                                                                                                                            521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          270
                                        g
                                                                                 Qγ
                                                                                                               g
                                                                                                                                                          Ω
                                                                                                                                                                                               qq
                                                                                                                                                                                                                                        Qγ
                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                   Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 43;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 4, Application US/09296284A
Patent No. 6204040
GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Choi, Eui-Sung
APPLICANT: Lee, Eun-Hae
TITLE OF INVENTION: and Methods of Use Thereof
FILE REFERENCE: 1533.0870000
CURRENT APPLICATION NUMBER: US/09/296, 284A
CURRENT FILING DATE: 1999-04-22
NUMBER OF SEQ ID NOS: 87
SOFTWARE: PatentIn Ver. 2.0
--VHAPKNGFF---'YIIDA-KTGKFISGKPYT-----YENWAN 359
                                                               360 GLDPVIGRPNYNP------DALWILLNGKPWYGIPGDLGGHNFAAAMAYSPQ---TK 405
                                                                                                                                                          406 LVYIPAQQVPFV-----YDPQKGGFKAH------HDSWNLGLDMNKIGLLDD 446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EGTPLIVDGVMYATTNWSKMKALDAATGKLL...----WSYDPKVPGNIADRGCCDT 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 76 INTNKVWAAVGMYTNSWDPNDGAI--LRSSDQGATWQITPLPFKLGGNMPGRGMGERLAV 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  142 VNRGAAYWNGKVYFGTF---DGRLIALDAKTGKLVWSVYTVP----KEAQLGHQRSYTV 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             134 DPNNDNILYFGAPS-GKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVW 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VAF---DKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDP 249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQH 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            284 VTDLVYLGVGN-GSPWN-----YKFR------SEGKGNNLFLGSIVAIN--- 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           194 D------GAPRIAKGKVIIGNGGAEFG-ARGF-------VTAYDAETGKMDW 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PNTIMVATQISWWPDTIIFRSTDGGATW-----TRIWDWTSYP------NRSLRYV 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      26 EGAP----GILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVV--SIA----ADP 75
                                          G-----GQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTS
                                                                                                                    PVFTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPG--GVTTG
                                                                                                                                                                                                 GTVAASADGSRFV------WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVN
                                                                                                                                                                                                                                        447 NDPQHKADKAQFLKDLKGWIVAWDP--QKQQAAFTVDHKGPW----------
                                                                                                                                                                                                                                                                              579 PKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEG-DLW-LAASSGLY
                                                                                                                                                                                                                                                                                                                     -----NGGLL-ATACGVLFQGLANG-----EFHAYDATTGKDLFTFPAQSAII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RFFTVPNPDNKPDGAASDDVLMSKAYPT----WGK--GGAWKQQGGG--GTVWDSLIYDP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               73; Mismatches 235; Indels 258;
                                                                                                                                                                                                                                                                                                                                                            ----HSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTG 682
                                                                                                                                                                                                                                                                                                                                                                                  APPUTYTANGKQ-----YVAVEVGW-----GGIYP--FFLGGVARTSG 564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 754;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 141; DB 4
Pred. No. 0.026;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TYPE: PRT Gluconobacter suboxydans US-09-296-284-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 3.5%;
Best Local Similarity 20.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches 145; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQ ID NO 4
LENGTH: 754
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-296-284-4
    327 L-
                                                                                                                    470
                                                                                                                                                                                                 528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 232
                                                                                                                                                                                                                                                                                                                     487
                                                                                                                                                                                                                                                                                                                                                          637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     321
                                                                             셤
                                                                                                                                                  g
                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                             οp
                                          ò
                                                                                                                    ö
                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

Methods of Use 41; ---VHAPKNGFF---YIIDA-KTGKFISGKPYT----YENWAN 393 528 GTVAASADGSRFV-----WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVN 578 PKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEG-DLW-LAASSGLY 636 ------NGGLL-ATAGGVLFQGLANG-------EFHAYDATTGKDLFTFPAQSAII 562 470 PVFTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPG--GVTTG 527 64 F---APPTRPVPWSTPLQATAYG-PACPQQF-----NYPEELREITMAWFNTPPPSAGES 114 STDGGAT---WT--RIWDWISYPNR-----SLRYVLDISAEPWLIFGVQPNPPVPSPKL 378 150 GLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSSLGQASKTI 209 Indels 309; Gaps 16 GLSASLASGAPTVKI----DAGMVVGTTTTVPGTTATVSEFLGVPF-----AASPTR 63 APPLICANT: Duvick, Jonathan P.
APPLICANT: Gilliam, Jacob T.
APPLICANT: Gilliam, Jacob T.
APPLICANT: Maddox, Joyce R.
TITLE OF INVENTION: Aniho Polyol Amine Oxidase
TITLE OF INVENTION: Polynucleotides and Related Polypeptides and Meth
FILE REFERENCE: 1134
CURRENT APPLICATION NUMBER: US/09/352,159A
CURRENT FILING DATE: 1999-07-12
EARLIER APPLICATION NUMBER: 60/192,936
EARLIER APPLICATION NUMBER: 60/195,391
EARLIER FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 46
NUMBER OF SEQ ID NOS: 46
SOFTWARE: FRASESO for Windows Version 3.0 GDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTIIFR 415 G-----GQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTS 440 LVYIPAQQVPFV-----YDPQKGGFKAH-----HDSWNLGLDMNKIGLLDD 210 FVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSS Length 1000; 637 ----HSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTG 682 563 APPVTYTANGKQ-----YVAVEVGW-----GGIYP--FFLGGVARTSG 598 181 NDPOHKADKAOFLKDLKGWIVAWDP--QKQQAAFTVDHKGPW-----

DD 16 GLSASLASGAPTVKIDAGMVVGTTTTVPGTTATVSEFLGVPFAASPTR 63  210 FVGVADPNNPVEMSEDGATWQAVGAPTGFIPHKGVEDPVNHYLYTATSNTGGPYDGSS 269	0y 621 PGKE
Qy         379 GWMDEAMAID	RESULT 13 1962-05 1968-05 1968-05 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25 1969-06-25

```
-----IKEGLTETTFEE-----DGV----ATGN-----HEYCVEVKYTAGVSPKECVNV 1375
                                      SVDYAELNP---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM: Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-352-159-29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-352-159-29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT
                                    477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         205
                                                                     g
                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                           d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            52;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 977 WKTIDADGDGNNWTTTPPP---GGT---SFAGHNSAICASSASYINFEGPONPDNYLVTP 1030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IQ----FTMGGSPTP----- 1335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1091 IRGTRVQGTWYQKTVQLPAGTKYVAFRHFGCTDFFWINLDDVEIKANGKRADFTET-FES 1149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1266 GLS-TEADGAKPQSV------W----IERTVDLPA-GTKYVAFRHYNCSDLNYILLDD 1311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----DPNNDNILY-- 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----TDTTGYQSDIQ---GVVWVAFD------KSSSSLGQASKTIFVG 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            213 VADPNNPVFWS---RDG-GATWQAVPGAPTGFIPHKG------VFDPVNHVL- 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----YIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDY-----F 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      299 GYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVL--D 356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -FGAPSGKGL--W-----VGTYIANP-- 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGT 476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            69 VSIAADPINTNKV----WAAVGM-YTNSWD-------PNDGAILRSSD----- 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          357 ISAEPWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 2; Length 2628; 0.33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          265;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----QGATWQITPLPFKLGGNMPGRGMGERLAV---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          68; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3.4%; Score 136; 20.1%; Pred. No. 0
                                                                                                                                                                                                        PRIOR APPLICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 372-5800
INFORMATION FOR SEQ. ID NO: 14:
SEQUENCE CHARACTERISTICS:
                                                                                                                                          PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
                                                                                        APPLICATION NUMBER: US/08/570,311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2628 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                     CURRENT APPLICATION DATA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; MOLECULE TYPE: protein US-08-570-311-14
                                                                                                       FILING DATE:
CLASSIFICATION: 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
Matches 178; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
Sequence 29, Application US/09352159A

Sequence 29, Application US/09352159A

Sequence 29, Application US/09352159A

Select No. 6211344

SERVET INFORMATION:

APPLICANT: Buddox, Jonathan P.

APPLICANT: Gilliam, Jacob T.

APPLICANT: Maddox, Joyce R.

TITLE OF INVENTION: Amino Polyol Amine Oxidase

TITLE OF INVENTION: Application Polyol Amine Oxidase

TITLE OF INVENTION: Polyol Amine Oxidase

TITLE OF INVENTION: Application Polyol Amine Oxidase

TITLE OF INVENTION: Application NUMBER: US/09/352,159A

CURRENT APPLICATION NUMBER: 60/092,936

EARLIER APPLICATION NUMBER: 60/092,936

EARLIER PILING DATE: 1999-07-12

SEALIER FILING DATE: 1999-05-21

NUMBER OF SEQ ID NOS: 46

SOFTWARE: FastSEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     49;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1554 GTWYQKTVQLPAGTKYVAFRHFGCTDFFWINLDDVEIKANGKRADFTETFESSTHGEAPA 1613
                                                                         1376 TVDPVQFNPVQNLTGSAVGQKVTLKWDAPNGTPNPNPNPNPGTTTLSESFENGIPASWKT 1435
                                                                                                                                                                                                                           1436 IDADGDGNNWTTTPPPGGTSFAGHNSAICASSASYINFEG-PQNPDNYLVTPELSLPNGG 1494
                                                                                                                                                                                                                                                                                                                                                                                                                 1495 TLTFWVCAQDANYASEHYAVYASSTGNDASNFANALLEEVLTAKTVVTAPEAIRGTRVQ- 1553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1614 EWTTIDADG------DGQGWLCLSSGQLGWLTAHGGTNVVASFSWNGMALNPDN 1661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       94 PNDGAILRSSD-----QGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAPS 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     187 PQIDKYLKSSKYIAWPLQG--WQAT---FG-GGDHP------PKSDLV----PR 224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GKGLWRST----DSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSSLGQ 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGP 264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          265 YDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPD 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        315 SAGESEDCLNLNIY-----WPGTENTN--------KAVMV----W-- 342
                                                                                                                                                                                                                                                                                                                                   ---WAASQGVPANAQIRSDRVNP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GSPEFAPTVKIDAGMV------VGTTTTVP-GTTATVSEFLGVPFA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   580 KTFY----ALSNGTFY----------RSTDGGVTFQPVAAG----
                                                                                                                                                               509 --TDG-GKNWFQGSEPGGVT-TGGTVAASADGSRFVWAPGDPGQP-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----LPSSGAVGVMFHAVPGKEGDLWLAASSGL--YHSTNGGS-----SWSA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels 349;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 1205;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1662 YLISKDVTGATKVKYYYAVNDGF----PGDHY-AVMISKTGTNAG 1701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  648 -----ITGVSS-----AVNVGFGKSAPGSSYPAVFV--VGTIGG 679
---SIIVR----AGSFDPSSQPNDRHVAFS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ouery Match 3.4%; Score 135.5; DB 4; Best Local Similarity 21.3%; Pred. No. 0.12; Matches 176; Conservative 71; Mismatches 230;
                                                                                                                                                                                                                                                                                                                                   --VVYAVGFGNS------
```

Qγ	325	TIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPPV	373
qq	343	IXGGALEYGWNSFHLYDGASFAANQDVIAVINYRTNILGFPAAPQLPI	391
ò	374		394
q	392	:	451
ολ	395		450
qq	452		503
Οy	451	LG-GFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAF	507
qq	504		537
Οy	208		558
qq	538	GIVANDGLLFVLGENDTQAYLEEAIPNQPDLYQTLLGA	575
Qy	. 559	SWAASQGVPANAQIRSDRVNPKTFYALSNGTF	290
q	576	YPIGSPGIGSPQDQIAAIETEVRFQCPSAIVAQDSRNRGIPSWRY-YYNATFENLELFPG	634
Qy	591		615
qq	635		693
QY	616		654
qq	694		749
οy	655		869
qa	750	VVVGAGLSGLETARKVQAGLSCLVLEAMDRVGGKTLSVQSGPGRTTINDLGAAWIND	807
δλ	669	DQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGD 733	
qq	808	SNOSEVSRLFERFHLEGEL-ORTTGNSIHOAODGTTTTA-PYGD 849	

Search completed: July 2, 2002, 09:12:54 Job time: 62 sec

THIS PAGE BLANK (USPTO)

probable PPE prote probable PPE prote hypothetical prote hypothetical prote mucin MUC5B, trach OmpA-related prote hypothetical prote mucin, submaxillar

prote prote

hypothetical hypothetical

hypothetical prote probable oxidoredu beta transducin-li

beta-amylase (EC 3 hypothetical 367K

Sequence:

on:

Run

Searched:

Database

Result No.

```
A;Cross-references: EMBL:AL031515; PIDN:CAA20642.1; GSPDB:GN00070; SCOEDB:SC5C7.30c A;Experimental source: strain A3(2) C;Genetics: C;Genetics: A;Gene: SCOEDB:SC5C7.30c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       probable secreted cellulase - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      66 NGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGR 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                424 VKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAEL 483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     B.G.; Rajandream, M.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKW-DSGGQIHIAPM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6 YTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  186 DIQGVVWVAFDKSS-SSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            58.7%; Score 2371; DB 2; Length 890; 57.7%; Pred. No. 1.3e-131;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A;Residues: 1-890 <SEE>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C, Accession: T35237
R; Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, Submitted to the EMBL Data Library, September 1998
A; Reference number: 221572
A; Accession: T35237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 57.7%; Pred. No. 1.3e
Matches 431; Conservative 109; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                  ALIGNMENTS
A55547
A70524
G83559
G83559
T34433
T45025
A87336
T03099
S77300
T35248
T45248
T45248
T45248
T45248
  809
2204
3157
4180
4180
1055
11821
13288
3016
348
1196
820
820
  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         305
  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                         ; Search time 54.67 Seconds
(without alignments)
1300.642 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hypothetical prote
tail-host specific
probable PPE prote
outer membrane sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         hypothetical prote
probable RTX famil
hypothetical prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    probable flagellin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hypothetical prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      nucleoporin - rat
probable structura
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hypothetical prote
amylase A-180 - al
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         large repetitive p
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               hypothetical prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       extracellular seri
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              hypothetical prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            proteoglycan core
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            probable secreted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  related to stress
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        hypothetical phypothetical phy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     probable PPE
hypothetical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hypothetical
hypothetical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        hypothetical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Description
                                                                                                                                                                                                                                                                                                         1 ATTQPYTWSNVAIGGGGFVD......YIGTNGRGIVYGDIGGAPSG
                      GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                283138 segs, 96089334 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                    protein search, using sw model
                                                                                                                                                                 09:13:57
                                                                                                                                                                                                                                                                                                                                                                                 Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             B85547
F90696
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AB3528
H87323
H90681
D70533
A44345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                T49741
T16580
S10789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          T35237
D97013
T00349
F72393
A83412
E95965
S75251
S76412
T13256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AE1905
AE2254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     E70969
H98323
AD0835
B86807
T34434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A28452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           length: 0
length: 2000000000
                                                                                                                                                                                                                                                          US-09-917-376-3
4036
                                                                                                                                                                 2, 2002,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2468
39174
4199
4199
1908
993
993
902
902
902
1468
903
13055
13055
1468
13055
1468
13055
1468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     pir1:*
pir2:*
pir3:*
pir4:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PIR_71:*
                                                                                                                                                                                                                                                                                                                                                          BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               oţ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2371
2009
160.5
20238
20238
20238
186.5
186.5
172
175.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.5
166.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         160
160
158.5
158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score
                                                                                                                                                                                                                                                                                        Perfect score:
                                                                                                                                                                                                                                                                                                                                                          Scoring table:
                                                                                                                    ı
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Total number
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Minimum DB
Maximum DB
                                                                                                                 OM protein
```

10;

Gaps

24;

Indels

185 160

220

280

337

397 423

244

us-09-917-376-3.rpr

OY 362 WLTFG-VQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATUDLTKWDSGGQIHI 420	Oy 718 RRVVIGTNGRGIVYGDIGGA 737	Query Match  Query Match  A1.6%; Score 1680; DB 2; Length 856; Best Local Similarity 46.2%; Pred. No. 4.8e-91; Matches 346; Conservative 113; Mismatches 258; Indels 32; Gaps Oy 1 ATTOPYTWSNVAI GGGGFVDGIVENEGAEGILYVRTDIGGAYRNDANGRWIPLLDWNG 59 1;   1   1   1   1   1   1   1   1   1
	RESULT 2 D97013 D97014 C; Pobably secreted sialidase, several ASP-boxes and dockerin domain [imported] - Clostrid C; Species: Clostridium acetobutylicum C; Species: Clostridium acetobutylicum C; Species: Clostridium acetobutylicum C; Accession: D97013 C; Accession: D97013 B; Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J. Barteriol. 183, 4823-4838, 2001 A; Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clc A; Reference number: A96900; MUID:21359325; PMID:21359325 A; Recence number: A96900; MUID:21359325; PMID:21359325 A; Residues: DNA A; Residues: 1-839 < KUR> A; Molecule type: DNA A; Residues: 1-839 < KUR> A; Residues: 1-839 < KUR> A; Residues: Clostridium acetobutylicum ATCC824 A; Experimental source: Clostridium acetobutylicum ATCC824 C; Genetics: A; Gene: CAC0919 Query Match Best Local Similarity 50.9%; Pred. No. 2.2e-110; Matches 377; Conservative 107; Mismatches 240; Indels 16; Gaps 12;	Qy         3 TQPYTWSNVAIGGGGFVDGIVFNEGAPGILIYVRTDIGGMYRWDAANGRWIPLLDWVGWNN 62           1

17;

Db 187SVTEVLEDERSGERGSPTRIFVGGSEPRG-IFVTEDGGTTWNVLPNLDLIP 239  Qy 243 HKGVFDPVNHVLXIATSNTGGPYDGSSGDVMKFSVTSGTWTRISPVPSTDTANDYFGYSG 302  1	477 SVDYAELNPSIIVRAGSFDPSQPNDRHVAFSTDGGKNWFQ-GSEPGGVTTGGTVAAS	QY 644 SWASHIGDANGERSAPESSYPAVENGINGARRANGIAN OF 114  Db 615 TERLCHVDIAYVIGERAPESTYPALTICH OF 1   1   1   1   1   1   1   1   1   1	RESULT 5 A83412 hypothetical protein PA1874 [imported] - Pseudomonas aeruginosa (strain PA01) C;Species: Pseudomonas aeruginosa C;Species: Pseudomonas aeruginosa C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000 C;Accession: A83412 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L.; Lory, S.; Olson, M.V. Nature 406, 959-964, 2000 A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa A;Reference number: A82950; MUID:20437337 A;Status: preliminary A;Status: preliminary A;Actesion: A83412 A;Status: preliminary	A; Residues: 1-2468 <sto> A; Residues: 1-2468 <sto> A; Cross-references: GB:AE004613; GB:AE004091; NID:g9947856; PIDN:AAG05263.1; GSPDB:GN A; Experimental source: strain PAO1 C; Genetics: A; Gene: PA1874  Query Match Query Match Best Local Similarity 22.8%; Pred. No. 9e-06; Matches 205; Conservative 88; Mismatches 314; Indels 292; Gaps 50; QY 11 VAIGGGGFVDGIVFTNGGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY 65                                      </sto></sto>
QY         296 DYFGYSGLTIDRQHPWTIMVATQISWWPDTIIFRSTDGGATWTRIMDWTSYPNRSLRYVL         355           b         312 TYGYGGLSVDLQVPGTLMYAALNCWWPDELIFRSTDGGATWSPINEWNGYPSINYXSX         371           QY         356 DISABEWLTFGVQPNP-PVPSFKLGWMDEAMAIDPFNSDRALYGTGATLYATNDLTKWDS         414           D         372 DISNAPWIGDTTSTDQFPVRVGWMVEALAIDPFDSNHMLYGTGLTVYGGHDLTNWDS         428           QY         415 GGQIHIAPMYKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVFSTIFTSPVFTT         474           QY         415 GGQIHIAPMYKGLEETAVNDLISPPSGAPLISALGDLGFTHADVTAVFSTIFTSPVFTT         474           DD         429 KHNYTVKSLAVGIEEMAVLGLITPPGGAPLLSAVGDGGFYHSDLDAPMQYTHTPTYGT         488           QY         475 GTSVDYABLNPSITVRAGSFDPSSQPNDRHYAFSTDGGRYNPGGSEPGGYTGGTVAASA         534           DD         489 TNGIDYAGNPSANDVANTARGSNPGSTWANDAYAASTGTGAVALSA         543           OV         535 DGSRFWWAPGDPGOPVVANAGFGNSWABARGANDANDATEVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFGANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSALSANGTFVALSANGTFVALSALSANGTFVALSANGTFVALSANGTFVALSANGTFVALSALSANGTFVALSANGTFVALSANGTFVALSANGTFVALSANGTFVALSANGTFVALSANGTFVALSANGTFVALSANGTFVALSANGTFVALSANGTFVALSANGTFVALSANGTFVALSANGTFVALSANGTFVALSANGTFVALSANGTFVALSANGTFVALS		718 GDLQTYGRVFRGHERPCHLLRQSQREPAG 746  SULT 4  2393  pothetical protein TM0305 - Thermotoga maritima (strain MSB8)  Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000	Ribert, M. N. E.; Clayton, K. A.; Gill, S. R.; Gwinn, M. L.; Dodson, R. J.; Haft, D. H.; Hickey Garrett, M. M.; Stewart, A. M.; Cotton, M. D.; Pratt, M. S.; Phillips, C. A.; Richardson, D.; C. M. Nature 399, 323-329, 1999 A. Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sec A; Reference number: A72200; MuID:99287316 A; Reference number: F7239 A; Status: preliminary A; Molecule type: DNA A; Status: preliminary A; Molecule type: DNA A; Status: Status: GB. AE000512; NID:94980799; PIDN:AAD35393.1; PID:949808C A; Experimental source: strain MSB8 C; Genetics: A; Genetics: A; Genetics:	Query Match  28.88; Score 1160.5; DB 2; Length 707;  Best Local Similarity 35.28; Pred. No. 1.2e-60;  Matches 264; Conservative 123; Mismatches 275; Indels 89; Gaps 26;  Qy 6 YTWSNVAICGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGNNNWGY 65  1

us-09-917-376-3.rpr

QY         116         FKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQWTNFPDVGTYIA         175           Db         1188         -DAGGNPIGOYTAD	Science 293, 668-672, 2001 A; Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau hebalit, P., D., Vandenbol, M.; Vorholter, F.J.; Waldner, S.; Wells, D.H.; Wong, K.; Yeh, A;Title: The composite genome of the legume symbiont Sinorhizoblum meliloti. A; Reference number: A96039; MUID:21368234; PMID:11474104 C; Genetics: Anontation C; Genetics: A, Gene: SMD21548 A; Gene: SMD21548
1275 PIGETTADGSGNWSFTPGTPLANGTVVNAVAQDPAGNTGPQGSTTVDAVAPNTPVVNPSN 252 HVLYIATSNTGGPYGSSGDVWKFSVT-SGTWTRISPVPSTDTANDYFGYS	Query Match 5.0%; Score 202.5; DB 2; Length 2174; Best Local Similarity 22.7%; Pred. No. 0.00093; Matches 176; Conservative 90; Mismatches 312; Indels 197; Gaps 36;
Db 1335 GNLLNGTAEPGSTVTLTGGGTADGSGNWSFTPGSQLPNGTVVNVTASDAAGNT 1394.  Qy 302 GLTIDRQHPNTINVATQISWWPDTIFRRFDGGGATWTRIWD 342	OY 12 AIGGGFVDGIVFNEGAPGILYVRT-DIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVS 70
Db 1395 SLPATTTVDSSLPSIPQVDPSNGSVISGTADAGNTIIITDGNGNPIGQVTADGSGNWS 1452  Qy 343 WTSYPNRSLRYVLDISAEPWLTFGVQPNPPVSPKLGWMDEAMAIDPFNS 392	Qy 71 IAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFK 117
Db 1453 FTPGIPLPDGTVVNVVARSPSNVDSAPAVITVDGVAPAAPVIDPSNG 1499  Qy 393 DRMLYGTGATLYATNDLTKMDSGQIHIAPMVKGLEETAVNDLISPPSGAPLI 445    1	QY 118 LGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNPPDVGTYIANP 177  ::
Qy 446 SALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPS 497	
1605 GQVTADGSGNWSFTPGTPLANGSVINALAQDAAGNNSSPTSATVDSLAPAAPV 526TGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAASGGVFANAQIRSDRVNP	271 DVWKFSVTSGFWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTIIFRS
Db 1658 IDPSNGSVIAGTAEAGATVILTDGNGNPIGQVTADGSGN-WSFTPGTP 1704  Qy 580 KTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGD 626	QY 331 TDGGATWTRIWDWTSYPNRSLRYVLDISABPWLIFGVQPNPPVPS 375
QY 627 LWLAASSGLYHS-TNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVT 681	Db 1333DDAHGIVAQSIGGGGGVGGAGSSTRAEKEHSHTIDLVVGGSGGVGGSGGE 1382  Qy 428 EETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY 480  Db 1383 VDLSVGTTLS-TSGARALGLVAQSIGGGGGIGGAGEADSIASLVVGGSGGGTIDG 1436
320	QY 481 AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGG 528  1437 GAVTVDLTSQSSITTQGIAAHGLVAQSIGGGGGGGGGAASGAPLSFTGNSPGSYGDGG 1493
RESULT 6 E955967  hypothetical glycine-rich protein [imported] - Sinorhizobium mellloti (strain 1021) maga C; Species: Sinorhizoblum mellloti C; Species: Sinorhizoblum mellloti C; Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001	529 1494 586
Explain, 1930. Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001 A;Title: The complete sequence of the 1,689-kb psymB megaplasmid from the N2-fixing endc A;Reference number: A95842; MUD:21396508; PMID:11481431 A;Accession: E95965 A;Status: preliminary A;Molecule type: DNA	DD 1544 KSGNVTVSLDAGRTIQASGKDSIGIFAGSDAGFUNGTVTGCSGDNGAGTWV 1603  OY 630 AASSGLXHSTNGASGKDSIGIFAGSDAGTDNNGTIDVTVNGTVTGCSGDNGAGTWV 1603  OY 630 AASSGLXHSTNGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTI-GGVTGA 683  :
A; Residues: 1-2174 < KURA A; Cross-references: GB:AL591985; PIDN:CAC49389.1; PID:g15140875; GSPDB:GN00167 A; Experimental source: strain 1021, megaplasmid pSymB R; Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.	RESULT 7 S75251 hypothetical protein s1r1028 - Synechocystis sp. (strain PCC 6803) C;Species: Synechocystis sp. A;Variety: PCC 6803

Qy         542 APGD	hypothetical protein.slr0408 - Synechocystis sp. (strain PCC 6803) C;Species: Synechocystis sp. A;Variety: PCC 6803 C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000 C;Accession: S76412 R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Mo, C, M.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada DNA Res. 3, 109-136, 1996 A;Title: Sequence analysis of the genome of the unicellular cyanobacterium S	S. A; Reference number: S74322; MUID:97061201 A; Accession: S76412 A; Status: preliminary A; Molecule type: DNA A; Residues: 1-4199 ** CKAN> A; Residues: ENBL:D90915; GB:AB001339; NID:91653604; PIDN:BAA18541.1; A; Note: the nucleotide sequence was submitted to the EMBL Data Library, June C; Superfamily: Synechocystis hypothetical protein slr0408	Ouery Match Best Local Similarity 19.2%; Pred. No. 0.019; Batt Local Similarity 19.2%; Pred. No. 0.019; Matches 198; Conservative 124; Mismatches 319; Indels 389; Gaps Qy 9 SNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDANGRWIPLL 55   1	Qy         56 DWVGWNN-WGYNGVVSIAADPINTNKVW-AAVGMYTNSWDPNDGA-ILRSSDQGA 107           b         2318 FTVSTNNGQSWNSEVQLPQSAQTIFPPAIAFFNNVLYLAYVDGNNGLNIITSQDQGQ 2374           QY         108 TWQITPLPFKLGGNMPG	QY 160 TWSQMT	QY         250 VNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPST	347 PNRSLRYVLDISAEPWLIFGV	Qy 414SGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHAD 458
C; Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000 C; Accession: S75251 R; Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996 A; Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis A; Reference number: S74322; MUID:97061201 A; Accession: S75251 A; Status: nucleic acid sequence not shown; translation not shown	A; Residues: 1-3972 cKAN> A; Cross-references: EMBL:D90904; GB:AB001339; NID:g1652225; PIDN:BAA17165.1; PID:g165224 A; Cross-references: EMBL:D90904; GB:AB001339; NID:g1652225; PIDN:BAA17165.1; PID:g165224 A; Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996 C; Superfamily: Synechocystis hypothetical protein s1r1028 Query Match 4.7%; Score 189; DB 2; Length 3972; Best Local Similarity 19.3%; Pred. No. 0.013; Matches 166; Conservative 93; Mismatches 276; Indels 326; Gaps 42;	13 IGGGGFVDGIVFNEGAPGILIVRTDIGGMYRMDAANGRWIPLLDW-	VLOGPNSTIAVYGNANGITPLFFALGONNFT-1	Db 2693 SFATLDGWLYQAFYGINERISTKDSYIYIQRSRDGVSWENLTQVVPLDSKNLP 2752  Qy 201 -SLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHK 244	Db 2812 YVFVKDASDUDILYSSSSNPGSSGGWDGTSTVLTFSDVNQATNFPL 2858  Qy 301 SGLTIDRQHPNDILYNATQISWWPDTILFSTDGGATWTRIWD 342    Db 2859 SATVVPGLDGDTLAVAFRSNNSPATWVGLLNSSDVTNWGGSAELTQVDA 2907  Qy 343	Db 2908 NSQVSLTVVDGTYYLFFTSSTEASASYAFSTDGLNWGDITLIPWDDGNL 2956  Qy 379 GWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIH	LAVIAPGYRNLIQFPILDYPAINNLGGVFIYYGESGISVNDPPDVVLAAPDDL TTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNW	DD 5128 EYTANSFEGLGNLANNGTNNSQNFNPYGEVTTGLPRSQAGISISGGADVNGDGFSDFALG 3187

```
.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, akeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3B:AB001339; NID:91653604; PIDN:BAA18541.1; PID:9165
as submitted to the EMBL Data Library, June 1996
thetical protein slr0408
                                                                                                                                                                                                                                                                                                                                                                                 genome of the unicellular cyanobacterium Synechocys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                54;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | |: | : | 3247
                                                                                                                                                                                                                                                                                 lsion 25-Apr-1997 #text_change 20-Jun-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                :| : | : : |: : |: CGTLSLFAANNSTSTVLQFYLNSSNEWIYANEIGS 2431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     : | : | | | : | | | SNQINWSS-PQVITNNISQSPPAIAF - 2543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BDNGNQVAWGNSTGVLNSFTNIDYPETQTTATGVDLN 2748
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RG-MGERLAVDPNNDNILYFGAPSGKGLWRSTDSGA 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NNPV-FWSRDGGATWQAVPGAPTGFI----PHKGVFDP 249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----VATQISWWPDTIIFRSTDGGATWTRIWDWTSY 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NKVW-AAVGMYTNSW-----DPNDGA-ILRSSDQGA 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WKFSVTSGTWTRISPVPST------DTAN 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /----- QPNPPVPSPKL 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3T-----413
-- PGOPVVYAVGFGNSWAASOG--- VPANAQIRSDR 576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 186.5; DB 2; Length 4199;
Pred. No. 0.019;
; Mismatches 319; Indels 389; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----FNEGAPGILYVRTDIGGMYRWDAANGRWIPLL 55
                                                                                                                                                                                                                      echocystis sp. (strain PCC 6803)
                                                                                                                                                                                                                                                                                                                                                                                                                        17061201
                                                                                                                      897
```

Db 2749 SDGIPEIAIGSDERKIAGQISTSGSFSLLPTPTTSSVINTLAAANQLENIGDFNGDGIAD 2808	
QY 459 VTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPS 497  :	
SQPNDRHVAFSTDGGKNW	Db 1077 GYAVAYMGTNGNNG-HDGFPGKDGTGIKTTTTTYAGSTSGTTP 1118 Oy 310 PNTIMVATQISWWPDTIIFRSTDGGATWTR-IWDWTSYPNRSLRYVLDISAEPWLFFGVQ 368
Db 2869 GQISTYQLSRIAQAGDVNGDGFDDLLISSPYTVDAENNQGGVFVVFGGDDWNNQPFDLGQ 2928	
2929 LRANQSQGSNPRGFAIDGSPNSQAGIALNGGGDINGDGFADFIIGAPGENNLQYNQQIVF	QY         369 PNPPVPSFKLGMMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGQIHIAPM 423           :   :   :   :   :   :   :   :   :   :
OY 545PDGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYAL 585    ::	424 VKGLEETAVNDLISP-PSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFT
QY 586 SNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGK 623	1181 -KGIRATAITYQASPNGTTAPTGTWSASVPPVAKGSFLWTFFTJWTYTDNITE 474 TGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWF    :    :
QY 624 EGDLWLAASSGLYHSTNGGSSW-SAITGVSSAVNVGEGKSAPGSSYPAVFVVGTIGGVTG 682	1232 TGYAVAXMGINGANGANGARDGIGAATITILIYAGGISGSLIFFNNGW- 517 QGSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVXAVG
QY 683 AYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTN 725    :	DD 1280 1323  QY 556 FGNSWAASQGVPANAQIRSDRVNPKTFYALSNGFFRSTDGGVFQPVAAG 606
Qy 726 GRGIVYGDIG 735  Db 3195 GVWTQFHDDG 3204	607LPSSGAVGVMFHAVPGKEGDLMLAASSGLYHSTNGGSS
RESULT 9 T13256 tail-host specificity protein homolog _ Lactococcus lactis phage BK5-T	Qy 645
C:Species: Lactococcus lactis phage BK5-T C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 13-Aug-1999 C:Accession: T13256 R;Boyce, J.D.; Davidson, B.E.; Hillier, A.J.	QY 683 720  QY 683 720
Appl. Environ. Microbiol. 61, 4089-4098, 1995 A.Title: Sequence analysis of the temperate Lactococcus lactis bacteriophage BK5-T and d A.Reference number: 217646; MUID:96064422 A.Accession: T13256	Qy 721YIGTNG 726               Db 1544 TGYAVAYMGTNG 1555
A.Status: preliminary; translated from GB/EMBL/DDBJ A.Molecule type: DNA A.Residues: 1-1904 <boy> A.Cross-references: EMBL:L44593; NID:9928826; PID:9928828; PIDN:AAA98579.1</boy>	RESULT 10 F70846
Query Match 4.5%; Score 182; DB 2; Length 1904; Best Local Similarity 20.7%; Pred. No. 0.012;	strain H37RV)  text_change 22-Oct-1999
Matches 189; Conservative 83; Mismatches 260; Indels 380; Gaps 55; Qy 1 ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGR 50	D.; Gor ; Holro S.
	Nature 393, 537-544, 1998 A; Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A; Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
OY 51 WIPLDWGWNNWGXNGVVSIAADPININNVWAAVGMYINSWDFNDGAILKSSDQGA 10/ 	A; Keterence number: A/USOU; MUID:9529596/ A; Accession: F70846 A; Status: preliminary; nucleic acid sequence not shown; translation not shown
QY 108 TWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATW-SQMTN 166 10	A; Molecule Lype: Una A; Residues: 1-2523 «COL> A; Cross-references: GB:AL021841; GB:AL123456; NID:g3261517; PIDN:CAA17115.1; PID:e125 A; Experimental source: strain H37Rv
167 FPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNN	C;Genetics: A;Gene: PPE
DD 9/4 VPAGGELWIKIVWIKTDN-ISEIGKSVAMMGVKGDKGDPGNNGIA 1020 Qy 219 PVFWSRDGGATWQAVPGAPIGFIPHKGVFDP 249	Query Match 4.3%; Score 175.5; DB 2; Length 2523; Best Local Similarity 21.8%; Pred. No. 0.044;

44;

```
A;Molecule type: DNA
A;Residues: 1-993 ckUR>
A;Residues: 1-993 ckUR>
A;Cross-references: GB:BA000019; PIDN:BAB72748.1; PID:g17130136; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr0791
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                745 -VYEIDSVGDV--IIEAADAGIDTVISSVD-----WTLGVNLENLTLVGNQATLGIGND 795
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           236 LLLPSIR-----PVGGLGSLPTSQGWLKFGASPPVFPLF---PVGGTQSLI 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            234 PGA---PTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPS 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       394 VT-------RYHLVVENNTYKLFAPDSSTPILSGNLRDYSAFNHSTAAPSP- 437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           :: ||: : |: | | |----ITSLPFDPYETPNFLFLGDNTTSAQASSNLTQVELQTNTRVRFVPNADYNGQANLT 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ---WD-----SGGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVP 463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        588 TGNDTLDGGEGSDLLYGNEDNDIINGGVGNDNLDG-----GTGDD----ILRGGTGNDIY 638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               639 TVDTVGDVIEENPNEGTDKVNSYISWTLGANLENLTLLGNTIIDG--TGNELDNHIIGNN 696
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         612 AVGVMFHAVPGKEGDLWLAASSG--LYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYP 669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     670 AVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNW-------GQAITGDHAN 716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | :|| | | : | :| :| | | :| 58 DNSNGTWEYTLNNGTSWTTFGTPSLTAARLLPSNANTKIRFHPNANFSGTADINFY--AW 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        280 TGGIQLNSSSIGSSG-YSNYNSYAPILFNQAFPELDPVKGFTISFDVKINGETHTSD-DN 337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        --IFRSTDGG 334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    338 GDGIQDRAGFSVIVVTSDKTKAI----ELGFWTDEIWAQTASPLFTHSTTERAFRNTTTA 393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    335 ATWIRIWDWISYPNRSLRYVLDISAEPWLIFGVQPNPPV---------PSPK 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        93 DPNDG------MPG------AILRSSDQGATWQITPL---PFKLGGN--MPG----- 124
                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----WAAVGMYTNSW 92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---VRAGSFDPSSQPNDRHVAFS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                176 ISNKGSLLADLVRGLISDSDPDYQGIAVTGADNNNGSWQYSLDGGVNWLNFGNASDNSAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IANPTDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            529 -SITVSPINNPIQGTTGLDKLYGTANEDIINGNEGNDYLFGRAGNDTLDGGEGNDYLFGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             T----DGGK--NWFQGSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSW-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---AASQGVPANAQIRSDRVNPKTFYA----LSNGTFYRST--DGGVTFQPVAAGLPSSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----VDPNNDNILYFGAPSGKGLWR-STDSGATWSQMTNFPDVGTY
                                                                                                                                                                                                                                                                                                                                                                                                            Indels 270;
                                                                                                                                                                                                                                                                                                                        Length 993;
                                                                                                                                                                                                                                                                                                                                                                                                        325;
                                                                                                                                                                                                                                                                                                                        ; DB 2;
0.02;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    378 LGWMDEAMAIDPFNSDRMLY-GTGAT-LYAINDLTK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            45 DAANGRW-IPLLDWVGWNNWGYNGVVSIAADPINTN-KV--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        291 IDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTI---
                                                                                                                                                                                                                                                                                                                                                                  ; Pred. No. 0.02
96; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STIFTSPV----FTTGTSVDYAELNPSII-----
                                                                                                                                                                                                                                                                                                                            Score 172;
                                                                                                                                                                                                                                                                                                                   4.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LRRVYIGTNGRGIVYGDIG 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LDNRITGNNADNVLFGEAG 814
                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 19.69
Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    125 -RGMGERLA-
                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     964
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A.Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120 (C)Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002 (C)Accession: AE1905 (C)Accession: AE19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               outer membrane secretion protein alr0791 [imported] - Anabaena sp. (strain PCC 7120)
    38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1866
                                                                                                                                      1522 NTGIGNSGIASTGLFNAGGFNTGVVNAGSYNTGSFNAGQANTGGFNPGSVNTGWLNTG-- 1579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   :||||: :: |: || :: |: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: || :: ||
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----GLFNSGTGNVGFFNSGTGN----WGVFNSGSY-NTGIGNSGIASTGLFNAGGF 2117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1615 LGFSYRPAVLPQ-----------TPFLDLTLTGGLGSVVIPAIDIPAIR 1652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1653 PEFSANVAIDSFTVPSIPIPQIDLAATTVSVGLGPITVPHLDIPRVPVTLNYLFGSQ--- 1709
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1710 -PGGPLKIGPITGLFNTPIGLTPLALSQIVIGASSSQGTITAFLANLPFSTPVVTIDEIP 1768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1867 GAMVSGGWNOAPSALLGGGSGVFNAGTL----HSGVLNFGSGMSGLFNTSVLGL----G 1917
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1958 -- GNVGLGNVGDFNLGAGNVGGFNVGGGNIGGNNVGLGNVGWGNFGLGNSGLTPGLMGLG 2015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2118 NTGVVNAGSYNTGSFNAGQAN - - TGGFNPGSVNTGWLNTGDINTGVANSGDVNTGAFISG 2175
                                                                                                                                                                                                                                                                                                                        -----DINTG-----GDYOGL 1614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1918 APALVSGLGSVGQ-----QLSGLLASGTA----LHQGLVLNFGLADVGL-- 1957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     657
                                                                                                                                                                                                                                                                                                                                                                                                        RGMGERLAVDPNNDNILLYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQ 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                233 VPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTD 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AP-LISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQP 500
                                                                           10 NVAIGGGGFVDGIVFNEGA--PGILYVRTDIGGMYRWDAAN-GRWIPLLDWVGWNNWGYN 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   599 TFQPVAAGLPSSGAVGV-MFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNV
                                                                                                                                                                                                                                                GVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSS-DQGATWQITPLPFKLGGNMPG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SDIQGVVWV-AFDKSSSSLGQ---ASKTIFVGVAD-----PNNPVFWSRDGGATWQA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            293 TANDYFGYS------GLTIDRQHPNTI-------MVATQISWWPDTIIFRSTDGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -AMAIDPFN-SDRMLYGTGATLYATNDLTKWDSGQIHIAPMVKGLEETAVNDLISPPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                501 NDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV-WAP---GD----PGQPVVY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            553 AVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTF-----YRSTDGGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 658 GFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQA----ITG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        335 ATWIRIWDWISYPNRSLRYVLDISAEPWLIFGVQPNPPVPSPKLGWMDE------
    Indels
327;
    Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   713 DHANLRRVYIGTNGRGIVYGDIG 735
92;
    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C; Species: Anabaena sp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Status: preliminary
175;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       11
                                                                                                                                                                                                                                                                                                                                                                                                        125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2176
                                                                                                                                                                                                                                                                                                                        1580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1828
    Matches
                                                                                                                                                                                                                                            29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    185
                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                        음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
```

587 260

411

substra

```
Excurations and insularly acceptance in Lampoileus) - Eschericina coli (Strain Old); Ho, Subbstra C; Species: Escherichia coli #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001 C; Accession: B85547
R; Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousls, K.; Apoda Nature 409, 529-533, 2001
A; Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A; Reference number: A65480; MUID:21074935; PMID:11206551
A; Accession: B85547
A; Molecule type: DNA
A; Molecule type: DNA
A; Molecule type: DNA
A; Molecule type: DNA
A; Cross-references: GB-AE005174; NID:g12513368; PIDN:AAG54838.1; GSPDB:GN00145; UWGP: A; Ceberimental source: strain O157:H7, substrain EDL933
C; Genefics: A; Generics:
                                                                                                                                                                                          probable RTX family exoprotein [imported] - Escherichia coli (strain O157:H7,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | :| :| | :| | 4314 VTMVGALNWTYSDSGLVSGAYTYSARVVDLAGNITSSSDF---VLTVDTSIPTTLAQITS 4370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4431 TVSATAYTVTAQVKSSAGNGNNANISNGTVTVNAAIDYTPTWTTASKTTAW---GLTYGL 4487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4270 ------QRQGTLSSLQ----ATDDTTPLLNGVLSAPLASGEVVYLYRNGLLGA 4313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --QISWW---PDT--- 325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              356 DISAEPWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSG 415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   416 GQIHIAPMVKGLEETAVNDLISPPSGAPLIS----ALGDLGGFT-HADVTAVPSTIFTSP 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  471 V-FTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFST--DGGKNWFQGSEPG---- 522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    27 GAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVG 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4110 ----WSYVDGRTLTNGT--TTWQVRVV--DLAGNV-GATSSQSALIDTVNPAQVLTIAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------DNILYF-----GAPSGKGLWRSTDSGATWSQMTNFPDVGTYIAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           177 PIDITG---YQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPN---NPVFWSRDGGATW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4371 QTTRDTTPIISGVITAALASGOYVEVVINGKTYTSEPGGAVVVDPAHNTWYVOLPDTDAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     326 -----IFRSTDGGATWTRIWDWTSYPNRSLRYVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4488 D-SHGMWIVLANQQVMQSTDP-LTWSKTALT------LYQSGNN-YATSSIADYDRN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDG-----SSGDV-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            273 ------WKFS---VTSGTWT---RISPVPSTDTANDYFGYSGLTIDRQHPNTIMVAT-
                                    676 TIGGVTGAYRSDDCGTTWVLINDDQHQYG--NW---GQAITGDHANLRRVYIGTNGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4083 GANEFAQISTDNGA------TWVN------VTVAADSLN----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         87 MYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNN-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 5188;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
4.2%; Score 170.5; DB 2;
Best Local Similarity 20.3%; Pred. No. 0.22;
Matches 191; Conservative 82; Mismatches 269;
                                                                                                                                       13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     138
                                                        746
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   231
     ŏ
                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Qγ
hypothetical protein alr3588 [imported] - Anabaena sp. (strain PCC 7120)
C.Species: Anabaena sp.
A.Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C.Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C.Accession: AE2254
B.Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, Shille: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana A; Reference number: AB1807; WUID:21595285; PMID:11799840
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            41;
                                                                                                                                                                                                                                                                                                                                                                                       A;Cross_references: GB:BA000019; PIDN:BAB75287.1; PID:g17132721; GSPDB:GN00179 A;Experimental source: strain PCC 7120 C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          172 GQ--VIDQND----YXS-----FTLTNSGTV-----DINLNSLNGTDT--LYADL 208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         245 GVFDPVNHVLYI------ATSNTGGPYD-------GSSGDVWKFSVT 278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          246 SQSDPGNYNLEFNFSADPPDAGGNTSDTSTPINLPATFSEIISDQVSLGDSSDYQFTLA 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              306 SASLVEIQFTSLTADANLY-----LQTQNGGNILSSTQPGTALDAVRLSLNAGTYNIL 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            359 VNRGSTETAQYTLSGFAQAIGNDQAPNSTTLALNLNLNSPISLNEFVGNIDTNDYYKFTV 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              : : | : : | : : | : : | 419 NGTTEIN------LDLSILNSYLLDPQLVNADVQILNSGGTQVAISNQTGNSNESINTIL 472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      474 TGTSVDYAEL---NPSIIVR---AGSFDPSSQ---PNDRHVAFSTD--GGKNW-FQG--- 518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     131 LSAEVIQONNI-------NDNTIANARNIGTENTITSITPVIDYVSGS----- 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        128 GERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDI 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                188 QGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVP---GAPTGFIPHK 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QPNPPVPSPKLGWMD----EAMAIDP--FNSDRMLYGTGATLYATNDLTKWDS----- 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           415 -GGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFT 473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQG---ATWQITPLPFKLGGNMPGRGM 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          209 QLI-----SSSNSVIQTSATV------GTSLESISRSLAAGTYYIRAY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---GATWTRIWDWTSY------PNRSL------RYVLDISAEPWLTFGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                473 GAGTYFIRVYTSGLANTFYDLNITAQSQALLVQDINPTGNSDPANLTTLGNTLY----FT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----SEPGGVTTGGT----VAASADGSRFVWA-PGDPGQPVVYAVGFGNSWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PTFTTTFNNQL-----FFTAQNNSQLWSTDGTIGGTQVISAGGITKSTPRNL---T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     620 VPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSA----VNVGFGKSAPGSSYPAVFVVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             562 ASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDG--GVTFQPVAAGLPSSGAVGVMFHA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       589 NRISDINVGAGNSNPGNLTVAGNKLFFTAVDNDSIRKLWVYNGTNVNLVDVNASFSNSST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1. 4.2%; Score 171; DB 2; Length 908; Similarity 20.7%; Pred. No. 0.021; Conservative 98; Mismatches 272; Indels 246;
                                                                                                                                                                                                                                                                                                                         A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-908 <KUR>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         161;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Gene: alr3588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local S
Matches 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    519
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

47;

Page 9

us-09-917-376-3.rpr

Db 4567 IQVTVGTLTWYGSI	= :: .
Oy 523GVTTGGTVAASADGSRFVWAPGDPGOPVVYAVGFGNSWAASOGVPANAOIRSDR 576	Db 4474 QTTRDTTPIISGVITAALASGQYVEVVINGKTYTSEPGGAVVVDPAHNTWYVQLPDTDAL 4533
4606 FLWNNAGTLVGNSTTSNSGGSATVGGAVTGYLSLNEGSGVDLNNDGRIDL	326IIFRSTDGGATWTRIWDWTSYPNRSLRYVL
QY 577 VNPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGK-EG 625	DD 4534 TVSATAYTVTAQVKSSAGNGNNANISNGTVTVNAALDYTPTWT1ASKTTAWGLIYGL 4590 Ov 356 DISAEDWITFGVODNDPVPSPKIGWWDEAWAIDPRNSDRWIYGTGATLYATNDLTKWDSG 415
4656 VQHTYNLNNYYTLSSLINQGNGTFVWGQNTTNTFLSGAGSGAMSSSVSWTWADFDGDGDM	4591 D-SHGMWTVLANQQVMQSTDP-LTWSKTALTLXQSGNN-YATSSIADYDRN
OY 626 DLWLAASSGLYHSTNG	416 GQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFT-HADVTAVPSTIFTSP
QY 642 -GSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCG 690	DD 4639 GFESSA 4003 OV 471 V-FTTGTSVDYABLNPSIIVRAGSFDPSSOPNDRHVAFSTDGGKNWFOGSEPG 522
4776	4670 IQVIVGTLIWYGSIVAFDKEGDGYLDFWIGDAGGPDSNT
QY 691 TTWVLINDOQHQYGNWGQAITGGHANLRRVYIGTNGRGIVYGD 733 :::   ::	QY 523GYTTGGTVAASADGSRFVWAPGDPGQPVYAVGFGNSWAASQGVPANAQIRSDR 576  DD 4709 FLWNNAGTLVGNSTTSNSGSATVGGAVTGYLSLNEGSGVDLNNDGRIDL 4758
RESULT 14 F90696 bunchheter	QY 577 VNPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGK-EG 625
C;Species: Escherichia coli C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001	Qy 626 DLWLAASSGLYHSTNG
C;Accession: Fyloby R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Gasawara, N.; Yasunada, T.; Kuhara, S.; Shiba, T.; Hartori, M.; Shinada, H.	.; Han, C.G. Db 4819 DLFLPASQGRANYGSLLFNTNGVLGCPVAVGATATTYASQFSLAVDWNHDGLMDIARIAQ 4878
DNA Res. 8, 11-22, 2001 A.Title: Complete genome sequence of enterohemorrhagic Escherichia coli 0157:H A.Reference number: A90679: MITD: 01157231. PMITD: 1758706	Oy 642 -GSSWSAITGVSSAVNVGFGKSAPGSSYPAVEVVGTIGGVTGAYRSDDCG 690 H
A:Accession: F90696 A:Status: preliminary	OV 691 TTWVLINDDOHOYGNWGOAITGDHANLRRVYIGTNGRGIVYGD 733
A; Molecule type: DNA A; Residues: 1-5291 <hay> A: Cross_references: GB:BA000007; PIDN: BAB33965.1; PID: 913360000; GSPDB: GN00154</hay>	Db 4931 SVFLSRNTNTVSYGTSLHLRITDPNGINVYYGN
A; Axperimental source: Straim Olof:n/, Substraim Kimb Ologyoz C; Genetics: A. Gene F. Frenka?	RESULT 15
אינענווע: בינאסטאינ	A20452 proteoglycan core protein precursor, cartilage - rat
Query Match 4.2%; Score 168.5; DB 2; Length 5291; Best Local Similarity 20.1%; Pred. No. 0.3; Matches 190; Conservative 82; Mismatches 270; Indels 401; Gaps 46	
27 GAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVG	K.) Sodoge, K.; Sadaki, M.; holigar, E.; hassell, J.R.; ramada, I. J. Bolol. Chem. 262, 17757-17767, 1987 A; Title: Complete primary structure of the rat cartilage proteoglycan core protein de
4186 GANEFAQISTDNGATWVNVTVAADSLN	A; Reference number: A92623; MUID:88087070 A; Accession: A92623.
QY B/MYINSWDPNDGALLKSSDQGATWQITPLPFKLGGNWPGRGMGERLAVDPNN 137          :      ;      ;      ;        ;	A; Molecule type: mRNA A; Residues: 1-2124 <doe> R; Doege, K; Sasaki, M; Horigan, E.; Hassell, J.R.; Yamada, Y.</doe>
QY         138	J. BIOI. CHEMIN. 203, 100404, 1900 A;Reference number: A3005 A;Contents: annotation; revision to residue 698 R:Doege, K : Fernander, P : Hassell, J.R : Sasaki, M :: Yamada, Y.
	J. Biól. Chem. 261, 8108-8111, 1986 A, Title: Partial cDNA sequence encoding a globular domain at the C terminus of the ra
231 QAVFGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDV	A; Molecule type: mRNA A; Molecule type: mRNA A; Residues: 1856-2124 <do2></do2>
DD 4376 GTLSSSQATDDTTPLLNGVLSAPLASGEVVYLYRNGLLLGA 4416	A.Cross.references: GB:M13518; NID:9206104; PIDN:AAA41836.1; PID:9206105 R.Yomane, P.J.; Christner, J.E.; Baker, J.R. R.Yomane, P.J.; Christner, J.E.; 18778. 1487
QY 273WKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVAT- 317    :   :     :     :	A; Title: Cartilage proteoglycan aggregates. The link protein and proteoglycan amino-t A; Reference number: A28453; MUID:88087071 A:Accession: A28453
	A; Residues: 20-37, W', 39-60, E', 62-64, X', 66-69; 70-83; 84, 89-148, L', 150-238, S', 240,

```
Search completed: July
Job time: 140 sec
C; Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG (5.Keywords: glycoprotein F):1-19/Domain: slyanal sequence #status predicted <SIG> F:10-19/Domain: slyanal sequence #status predicted <SIG> F:20-2124/Product: proteoglycan core protein #status predicted <MAT> F:20-2124/Product: proteoglycan core protein #status predicted <MAT> F:44-135/Domain: link protein repeat homology <LINK1> F:504-581/Domain: link protein repeat homology <LINK2> F:504-581/Domain: link protein repeat homology <LINK3> F:504-581/Domain: link protein repeat homology <LINK3> F:502-683/Domain: link protein repeat homology <LINK4> F:1014-2034/Domain: C-type lectin homology <LICH> F:1014-2034/Domain: Complement factor H repeat homology <FRD> F:2041-2097/Domain: Complement factor H repeat homology <FRD> F:2041-2097/Domain: Complement factor H repeat homology <FRD> F:2041-2097/Domain: Complement factor H repeat homology <FRD> F:205-239,333,387,611,667,1842/Binding site: carbohydrate (Asn) (covalent) #status predi
                                                                                                                                                                                                                                                                                                                                  44;
                                                                                                                                                                                                                                                                                                                                                                                             1038 SGIEDITVLPTGRENLETSASGVEDVSGLPSGKEGLETSASGI--EDISVFPTEAEGLET 1095
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1211 VGSASGALDFGKLPSGTLGSGQTPEASGLPS-GFSGEYSGVDIGSGPSSGLPDFSGLPSG 1269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1426 SGQPSGFPVLSGTTPGTPDLASGAMSGSGDSSGTTFVDTSLIEVTPTTFREEEGLGSVEL 1485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SSLSSGAFDGSGLVSGFPTVSLVDRTLVESITLAPTAQEAGEGPSSILEFSGAHSGTPDI 1604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1486 SGLPSGETDLSGTSGMVDVSGQSSGAIDSSGLISPTPEF-SGLPSGVAEVSGEVSGVETG 1544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          487
                                                                                                                                                                                                                                                                                                                                                                      NGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGR 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AGLPSSGA-----VGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAV-N 656
                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   --VGTYIAN-PTDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                227 GATWQAVPGAPTG-----FIPHKGVFDPVNHVLY----IATSNTG------GPYD-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -GSSGDVWKF----SVTSGTWTRISPVPSTDTANDYFG------YSGL---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --TIDRQHPNTIMVAT-----GGATWTRI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            341 WDWTSYPNRSLRYVLDISAEPWLTFGVQPNP---------PVPSPKLGWM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DEAMAIDPFNSD-----RMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEETAVNDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----NPKTFYALSNGTFYRSTDGGVTFQP---VA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1330 SGQTSGS-----LDVSGETSGFFDVSGQPFGSSGTGEGTSGIPEVSGQAVRSPDTTEI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----IETSPUFTTGTSVDYAELNPSI-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -IVRAGSFD-----PSSQPNDRHVAFS-----TDGGKN-----WFQGSEPGGVTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGTVAASADGSRFVWAPG-----DPGQPVVYAVGFGNSWAAS------QGVP
                                                                                                                                                                                                                                                                                                                                  322;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    GMGERLAVDPNNDNIL-----YFGAPSGKGLWRSTDSGATWSQMTNFPD----
                                                                                                                                                                                                                                                                                      Query Match
4.2%; Score 168; DB 2; Length 2124;
Best Local Similarity 20.5%; Pred. No. 0.096;
Matches 190; Conservative 102; Mismatches 311; Indels 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ISPPSGAPLISAL-------GDLGGFTHADVTAV---PST-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ANAQIRSDRV-
                                                                                                                                                                                                                                                                                                                                                                      99
                                                                                                                                                                                                                                                                                                                                                                                                            979
                                                                                                                                                                                                                                                                                                                                                                                                                                                    126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1096
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1723
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   S G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 oy
D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             0.7
Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

2, 2002, 09:14:12

THIS PAGE BLANK (USPTO)

sapien

```
SWISS-PROT entry is copyright. It is produced through a collaboration sen the Swiss Institute of Bioinformatics and the EMBL outstation -
                       bacillus su
bacteriopha
                                                rattus norv
                                                          homo sapien
                                                                      dictyosteli
                                                                                  cellulomona
                                                                                            bacteriopha
                                                                                                         streptococc
            epstein-bar
                                                                                                                  caulobacter
                                                                                                                                 synechocyst
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -!- PTW: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE CHAINS, N-LINKED AND O-LINKED LOGOSACCHARIDES.
-!- SIMILARITY: CONTAINS 1 INMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
-!- SIMILARITY: CONTAINS 4 LINK DOMAINS.
-!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
-!- SIMILARITY: CONTAINS 1 SOSHI (SCR) DOMAIN.
-!- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
-!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIMILARITY).

-!- DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO TERMINIS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3, MAKES UP THE COOH TERMINIS. G1 CONTAINS LINK DOMAINS AND THUS CONSISTS OF THREE DISULEIDE-BONDED LOOP STRUCTURES DESIGNATED AS THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS) AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=86550698; PubMed=244893;
MEDLINE=86550698; PubMed=244893;
MEDLINE=86550698; PubMed=244893;
Doege K., Fernandez P., Hassell J.R., Sasaki M., Yamada Y.;
Partial cDNA sequence encoding a globular domain at the C terminus of the rat cartilage proteoglycan.";
J. Biol. Chem. 261:8108-8111(1986).

1- FUNCTION: THIS PROTEOGIVCAN IS A MAJOR COMPONENT OF EXTRACELLUIAR MATRIX OF CARTILAGENOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEOGIVCAN IN CARTILAGE. IT BINDS AVIDLY TO HYALDRONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION. MAY PLAY A REGULATORY ROLE IN THE MATRIX ASSEMBLY OF THE CARTILAGE.

1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLUIAR MATRIX (BY
                                                                                                                                                                                                                                                  ol-Kbb-1991 (Rel. 17, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
Aggrecan core protein precursor (Cartilage-specific proteoglycan core protein) (CSPCP).
                                                                                                                                                                                                                                                                                                                     Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-88087070; PubMed-3693370;
Dogge K., Sasaki M., Horigan E., Hassell J.R., Yamada Y.;
"Complete primary structure of the rat cartilage proteoglycan upprotein deduced from CDNA clones.";
J. Biol. Chem. 262:17757-17767(1987).
                      P16397
P03744
P52591
P35658
P22698
P50899
P07067
P29336
P35828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           gge K., Sasaki M., Horigan E., Hassell J.R., Yamada Y.;
Biol. Chem. 263:10040-10040(1988).
                                                                                                                                                                                                                            PRT; 2124 AA
                                                                                                                                                                 ALIGNMENTS
                                                         N214_HUMAN
SPG7_DICDI
GUXB_CELFI
NU98_HUMAN
VGP3_EBVA8
SUBF_BACSU
VG37_BPT4
N121_RAT
                                                                                                        GTFS_STRDO
SLAP_CAUCR
                                                                                                                                THIG SYNY?
                                                                                             VG37_BPT2
                                                                                                                                                                                                                            STANDARD;
920
886
11433
11026
11199
2090
532
1090
1341
1365
1005
656
                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                 01-AUG-1988 (Rel.
NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REVISION TO 698
                                                                                                                                                                                                                                                                                                              AGC1 OR AGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AND G3
133
132
131.5
131.5
131.5
131.5
131.5
130.5
130.5
130.5
                                                                                                                                                                                                                            PGCA_RAT P07897;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWI
between
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Doege
                                                                                                                                                                                                              PGCA_RAT
 rch time 28.05 Seconds
(without alignments)
1021.478 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          candida alb rattus norv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          micromonosp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            cellulomona
canis famil
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            clostridium
clostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 acinetobact
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sus scrofa
paenibacill
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            xanthomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     cellulomona
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gluconobact
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       porphyromon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  rabbit hemo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   gallus gall
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           synechocyst
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   rattus norv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 xanthomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           cellulomona
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        bacteriopha
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           vibrio chol
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       salmonella
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              vibrio sp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     petromyzon
                                                                                                                                                                                                                                                                                                                                                                  Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                      740
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Description
                                                                                                                                                    1 ATTOPYTWSNVAIGGGGFVD......YIGTNGRGIVYGDIGGAPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q59086
P12021
P21543
Q61282
P10342
P26501
P25927
P07898
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   P50400
P49793
P48839
Q02834
O05542
P75780
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q99715
Q60106
P50401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P27410
P08696
P29767
P02674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q9xd78
P73069
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P16112
P37198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             P07897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    P49791
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      028343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2990b5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         P37060
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     251845
                                                                                                                                                                                                                                      105224
           GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd
                                                                                                                                                                                                                                     hits satisfying chosen parameters:
                                                                               2, 2002, 09:17:53 ; Search
                                                                                                                                                                                                             105224 seqs, 38719550 residues
                                                                                                                                                                                                                                                                                                                                                                                                                    SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ALS3_CANAL
N153_RAT
QUIA_ACICA
                                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QUIA_XANCJ
Y48L_SYNY3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            XANP_XANS2
GUXA_CELFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PGCA_HUMAN
NU62_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AMYB_PAEPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GCA_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SOA_PSESP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUND_CELFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PGCA_CANFA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CA1C_HUMAN
                                                         OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOA_PSEAY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PGCA_CHICK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AGAA_VIBS7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    OHET_GLUOX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VG37_BPAR1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NANH_VIBCH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HAGA_PORGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NANH_CLOSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FIB1 PETMA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BIGA_SALT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NANH_MICV]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              YBIL_ECOLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GUNB_CELF1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POLN_RHDV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NU98_RAT
                                                                                                                                                                                      Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                            seq length: 0 seq length: 2000000000
                                                                                                                             US-09-917-376-3
4036
                                                                                                                                                                                                                                                                                                                                                                                                                                         Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                              SwissProt_40:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2124
11119
1468
809
809
11150
11196
2132
776
776
1953
347
747
790
347
937
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     647
757
760
1045
2333
3063
827
872
11103
2415
522
781
                                                                                                                                                                            BLOSUM62
                                                                                  July
                                                                                                                                                                                                                                      oę
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            167.5
167.5
166
157
153
152
149
148.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                145.5
145.5
145
145
143.5
142.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 142
142
139.5
139.5
139.5
138.5
137.5
137.5
137.5
136.5
136.5
134.5
134.5
134.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score
                                                                                                                                           Perfect score:
                                                                                                                                                                            Scoring table:
                                                                                                                                                                                                                                      Total number
                                                                                                                                                                                                                                                           Minimum DB :
Maximum DB :
                                                                                                                                                                                                                                                                                                                                                Database :
                                                                                                                                                       Sednence:
                                                                                                                                                                                                              Searched:
                                                                                  Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
```

```
the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                 SMART; SM00032; CCP: 1.
SMART; SM00034; CLECT; 1.
SMART; SM00445; LINK: 4.
PROSITE; PS01241; LINK; 4.
PROSITE; PS50041; C_TYPE_LECTIN_1; 1.
Glycoprotein; Carrelage; Proteoglycan; Lectin; Signal; Sushi; Repeat; Immunoglobulin domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              :NKED (GLCNAC. . .) (POT
E30BBE61593A34B1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC...
                                                                                                                                                                                                                                                                                                     AGGRECAN CORE PROTEIN IG-LIKE V-TYPE DOMAIN
                                                                                                                                                                                                                                                                                                                                                          C-TYPE LECTIN.
SUSHI.
G1-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIMILARITY.
SIMILARITY
                                                                                                                                                                                                                                                                                             POTENTIAL.
                                                                                                InterPro; IPR003066; Ig_MHC.
InterPro; IPR003565; Ig_V.
InterPro; IPR003348; Link.
InterPro; IPR003344; SGXXSG.
InterPro; IPR001344; SGXXSG.
InterPro; IPR001344; Iectin_c.
Pfam; Pr00159; SGXXSG; 58.
Pfam; Pr00239; SGXXSG; 58.
Pfam; Pr00193; Xllnk; 4.
ProDom; P0000918; Link; 4.
                                                                                                                                                                                                                                                                                                                        LINK 1
                                                                                                                                                                                                                                                                                                                                                                                                                                CS-1.
CS-2.
G3.
                                                                                                                                                                                                                                                                                                                                                                                                               G2-B'
                                                      EMBL; M13518; AAA41836.1; -.
EMBL; J03485; AAA21000.1; ALT_SEQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                           BY
BY
BY
BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     221117
                                                                                                                                                                                                                                                                                                                                                                                                                               1226
1909
2124
                                                                        PIR; A23835; A23835.
PIR; A28452; A28452.
HSSP; P98066; 1TSG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ą;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2041
2070
126
239
333
387
611
667
1842
                                                                                                                                                                                                                                                                                                                      170
268
504
602
1910
2040
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1914
1942
2010
                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DISULFID
                                                                                                                                                                                                                                                                                            SIGNAL
CHAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                       DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                            OMAIN
                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                               OMAIN
```

```
44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | | : | : | : | : | : | | : | | : | | | : | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | : | | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | : | | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1153 SASGVDVTGYPSGREDTETSVP--GVGDDLSGLPSGQEGLETSASGAEDLGGLPSGKEDL 1210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1426 SGQPSGFPVLSGTTPGTPDLASGAMSGSGDSSGITFVDTSLIEVTPTTFREEEGLGSVEL 1485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1486 SGLPSGETDLSGTSGMVDVSGQSSGAIDSSGLISPTPEF-SGLPSGVAEVSGEVSGVETG 1544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .663 EVTLITSELVEGVTEPTVSQELGHGPSMTYTPRLFEASGEASASGDLGGPVTIFPGSGVE 1722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1038 SGIEDITVLPTGRENLETSASGVEDVSGLPSGKEGLETSASGI--EDISVFPTEAEGLET 1095
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               304 --TIDRQHPNTIMVAT-------QISWWPDTIIFRSTD--GGATWTRI 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 487
                                                                                                                                                                          66 NGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGR 125
                                                                                                                                                                                                                                                                                                                                                    126 GMGERLAVDPNNDNIL------18GAPSGKGLWRSTDSGATWSQMTNFPD----- 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -GSSGDVWKF----SVTSGTWTRISPVPSTDTANDYFG-----YSGL--- 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---PVPSPKLGWM 381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      382 DEAMAIDPFNSD-----RMLYGTGATLYATNDLTKWDSGQIHIAPMVKGLEETAVNDL 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            527 GGTVAASADGSRFVWAPG-----DPGQPVVYAVGFGNSWAAS-------QGVP 567
                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  170 --VGTYIAN-PIDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   227 GATWQAVPGAPTG-----FIPHKGVFDPVNHVLY----IATSNTG------GPYD-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1330 SGQTSGS-----LDVSGETSGFFDVSGQPFGSSGTGEGTSGIPEVSGQAVRSPDTTEI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 466 -----IFTSPVFTTGTSVDYAELNPSI-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  568 ANAQIRSDRV-----VA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            605 AGLPSSGA-----VGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAV-N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1723 ASVPEGSSDPSAYPEAGVGVSAAP-----EASSQL----SEFPDLHGITSASRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      488 -IVRAGSFD------PSSQPNDRHVAFS------TDGGKN-----WFQGSEPGGVTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             657 VGFGKSAPG---SSYPAVFVVGT-----IGGVTGAYRSDDCGTTWV----LINDDQ
                                                                                         322;
4.2%; Score 168; DB 1; Length 2124; 20.5%; Pred. No. 0.043;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                436 ISPPSGAPLISAL-------GDLGGFTHADVTAV---PST-----
                                                                                         Indels
                                            Best Local Similarity 20.5%; Pred. No. 0.043; Matches 190; Conservative 102; Mismatches 311;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1119 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    341 WDWTSYPNRSLRYVLDISAEPWLTFGVQPNP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1828 TEISGEW----SDHTSEVNVTVST 1847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       701 HQY-GNWGQAITGDHANLRRVYIGT 724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ALS3_CANAL
        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ALS3_CANAL
ID ALS3_CA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

-----SATDVNSYTLSYANEYTCAGGYWQRAPFTLRWT 313

|: || | | || TYKNVPAGYRPEVDAYI

264

GWNN--WGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQI-TPLP

3

```
g
                                                                                                                                                                                                a
                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                           pp
                                                                                                                                                                                                                                                                                                                                                    Db
                                                                                                                                                                                                                                                                                                                                                                                         g
                            ò
                                            g
                                                                 à
                                                                                 qq
                                                                                                     ò
                                                                                                                                          ò
                                                                                                                                                                             οy
                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                 the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                          between the Swiss Institute of Bioinformatics and the EMBL outstation
                                                                                                                               Hoyer L.L., Payne T.L., Bell M., Myers A.M., Scherer S.; "Candida albicans ALS3 and insights into the nature of the ALS gene
                                                                Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Candida.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                          Curr. Genet. 33:451-459(1998).
-!- FUNCTION: MAY PLAX A ROLE IN ADHESION AND PATHOGENESIS.
-!- PTM: N-GLYCOSYLATED AND O-GLYCOSYLATED (FOTENTIAL).
                                                                                                                                                                                                                                                                                                      AGGLUTININ-LIKE PROTEIN 3.
10 X 36 AA TANDEM REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6A3FB3FC8C879A71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                   Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                  Repeat; Signal. POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                            POLY-THR.
POLY-THR.
POLY-THR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        POLY-THR.
POLY-THR.
POLY-THR.
POLY-THR.
POLY-THR.
                                                                                                                                                                                                                                                                                                                                                                                                                    POLY-THR
       15-JUL-1999 (Rel. 38, Created)
15-JUL-1999 (Rel. 38, Last sequence
15-JUL-1999 (Rel. 38, Last amoncatik
Agglutinin-like protein 3 precursor
                                                                                                            STRAIN=1161;
MEDLINE=98309840; PubMed=9644209;
                                                                                                                                                                                                                                                                                                                         1-1.
1-2.
1-2.
1-5.
1-6.
                                                                                                                                                                                                                                                                                   Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        119927
                                                                                                                                                                                                                                                                         EMBL; U87956; AAC39486.1;
                                                      Candida albicans (Yeast).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               723
759
845
987
1050
1061
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AA:
                                                                                                    SEQUENCE FROM N.A.
                                                                                 NCBI_TaxID=5476;
                                                                                                                                                                                                                                                                                  adhesion;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                     CHAIN
DOMAIN
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
REPEAT
                                                                                                                                                                                                                                                                                             SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                         REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                   Cell
```

```
646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              806
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     647 AITGVSSAV-----NVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQ 700
                                                                                                  364 TTTITTSYVGVTTSYSTKTAPIGETATVIV---DIPYHTTTTVTSKWTGTITSTTHTNP 420
                                                                                                                                                                                                         421 TDSIDTVIVQVPSPNPTVTTTEYWSQ--SFATTTTITGPPGNTDTVLIREPPNHTVTTTE 478
                                                                                                                                                                                                                                                             183 YQSDIQGVVWVAFDKSSSLGQASKTIFVGVADPNNPV----FWSRD--GGATWQAVPG 235
                                                                                                                                                                                                                                                                                                                                                                  --SSGDV 272
                                                                                                                                                                                                                                                                                                                                                                                                                   587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WKFSVTSGTWTRISPVPSTD-----TDNOYFGYSGLT------ID 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    700 FATITIVIAPPGGIDIVIIREPPNHIVITIEY-WSQSYATITIITAPPGEIDIVLIREPP 758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ------WDSGGQIHIAPMVKG 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           532 ASADGSRFVWAPGDPGQP----VVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALS 586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WSQSYTT-TTTVIAPPGGTDSVIIREPPNPTVTTTEYWSQSYATTTTITAPPGETDTVLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    476 TSVD----YAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   587 NGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWS
                                                                                                                                                                                                                                                                                                                                                                                                  GTDSVI----IKEPPNHTVTTTEYWSQSYTTTTTTTAPAPPGGTDTVLVREPPNHTVTTTEY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----PDTIIFRSTDGGATWTRIWDWTSYPNRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPPNHTVTTTE--YWSQSYATTTTTTAPPGETDTVLIREPPNHTVTT----TEYWSQS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           759 NHTVTTTEYWSQSYATTTTIIAPPGETDTVLIREPPNPTVTTTEYWSQ------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                807 -SYTTATTVTAPPGGTDTVIIYDTMSSSEISSFSRPHYTNHTTLWSTTWVIETKTITETS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        866 CEGDKGCSWVSVSTRIVTIPNNIETPMVTN-----TVDSTTTESTSQSPSGIFSESGVS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               920 VETESSTVTTAQTNPSVPTTESEVVF-----TTKGNNENGPYESPSTNVKSSMD-E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
01-MAR-2002 (Rel. 41, Last annotation update)
Nuclear pore complex protein Nupl53 (Nucleoporin Nupl53) (153 kDanucleoporin).
                                                                                                                                                        ----TTG
                                                                                                                                                                                                                                                                                           LEETAVNDLISPPSGAPLI - - - - SALGDLGGFT - - - - - HADVTAVPSTIFTSPVFTTG
GYRNSDAGSNGIVIVATTRTVTDSTTAVT---TLPFDPN-----RDKTKTIEILKPIP
                                                   -----FKLGGNMPGRGMGERLAVDPNNDN1LYFGAPSGKGLWRSTDSGAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LRYVLDISAEPWLTFGV----QPNPPVPSPKLGWMDE-----AMAIDPFNSDRML----
                                                                                                                                                                                                                                                                                                                                                                  236 APTGFIPHKGVFDPVNHVL-----YIATSNTGGPYDG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT; 1468 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ROHPNTIMVATOISWW-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1067 SOYNSDTOOTT 1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             701 HQYGNWGQAIT 711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N153_RAT
P49791;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  647
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      351
  314
                                                                                                                                                        161
                                                                                                                                                                                                                                                                                                                                                                                                                   532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
N153_RAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

42;

Gaps

303;

Indels

98; Mismatches 340;

Conservative

Local Similarity Les 170; Conserv

Best Loca Matches

Query Match

Length 1119;

DB 1;

4.2%; Score 167.5; DB 18.7%; Pred. No. 0.022;

us-09-917-376-3.rsp

-- 1136

```
----ESDQLAKATFAFGNQTNTTT--DQ 1176
                                                                                                                      1385 SSVFQFGSSTINFNFTNNNPSGVFTFGASPSTPAAAAQPSGSGGFSFSQSPASFTVGSNG 1444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PATHWAY: QUINIC ACID CATABOLIC PATHWAY, FIRST STEP. THIS PATHWAY ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO PROTOCATECHUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-96011389; PubMed-7592351;
Elsemore D.A., Ornston L.N.;
"Unusual ancestry of dehydratases associated with quinate catabolism
                                                                                                                                                                                                                                                                                             471 VFTT--GTSVDYAELNPSIIVRAGSFDPSSQPNDRHVA----FSTDGGKNWFQGSEPGGV
                                                                                                                                                                                                                    525 TTGGTVAASADGSRFVWAPGDPGQPVVYAVG-----FGNSWAASQGVPANAQIRSDRVN
                                                                                                                                                                                                                                               1283 TFGATTTSSSSGSFFVFGTGHSAPSASPAFGANQTPTFGQSQGASQ-----PN
                                                                                                                                                                                                                                                                                                                                   639 INGGSSWSAITGV----SSAVNVGFGKSAPG------SSYPAVFVVGTIG
                                           LDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDS
                                                                                                    415 GGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDL-GGFTHADVTAVPSTIF---TSP
                                                                                                                                                                                                                                                                           PKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLMLAASSGLYHS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               of Acinetobacter calcoaceticus for quinate-shikimate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
-!- INDUCTION: BY PROTOCATECHUATE.
-!- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             in Acinetobacter calcoaceticus ";
J. Bacteriol. 177:5971-5976 (1995).
J. Bacteriol. 177:5971-5976 (1995).
J. FUNCTION: CAN ACT EITHER ON QUINATE OR ON SHIKIMATE.
J. CATALYTIC ACTIVITY: Quinate + pyrroloquinoline-quinone = dehydroquinate + reduced pyrroloqquinoline-quinone.
J. CATALYTIC ACTIVITY: SHIKIMATE + PYRROLOQUINOLINE-QUINONE.
DEHYDROSHIKIMATE + REDUCED PYRROLOQUINOLINE-QUINONE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Quinate/shikimate dehydrogenase [Pyrrologuinoline-quinone] (EC 1.1.99.25) (NAD(P)-independent quinate dehydrogenase).
   16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                   Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FROM N.A., AND CHARACTERIZATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "The pca-pob supraoperonic cluster contains quiA, the structural gene dehydrogenase.";
                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Bacteriol. 176:7659-7666(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-95095936; PubMed-8002591;
                                                                        1137 -ESSSKPTFSFSVAK----PSVK---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Elsemore D.A., Ornston L.N.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Acinetobacter calcoaceticus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 1-20 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-BD413 / ADP1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -!- COFACTOR: PQQ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=471;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Acinetobacter.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PATHWAY
                                                                                                                                                                                                                                                                                                                                                                                                                                   QUIA_ACICA
059086;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                           355
                                                                                                                                                                                                                                                                             579
                                                                                                                                                                                                                                                                                                                                                                                                                      QUIA_ACICA
                                           ò
                                                                      g
                                                                                                    à
                                                                                                                              g
                                                                                                                                                            ò
                                                                                                                                                                                       셤
                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   31;
                                                                              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                192 WVAFDKSSSSLGQASKTIF---VGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFD 248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          884 GTSSSLNPAPSAFKFGIPSSSSGLSQTFTSTGN----FKFFGDQGGFKLGTSSDSGSTNTM 939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QHPNTIMVATQISWWPDTIIF-RSTD------GGATWTRIWDWTSYPNRSLRYV 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GATWQITPLP - FKLGGNMPGRCMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQ 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ---VGTYIANPTDTTGYQSDIQGVV 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          249 PVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   259; Indels 186;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Repeat; Zinc-finger; DNA-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 1468;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C3DFD9697C556A7C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4.1%; Score 166; DB 1
21.8%; Pred. No. 0.036;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RANBP2-TYPE 1.
RANBP2-TYPE 2.
RANBP2-TYPE 3.
RANBP2-TYPE 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; Pred. No. 0.03
71; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        164 MTNF-----------------------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLY-RICH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; L06821; -; NOT_ANNOTATED_CDS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROSITE; PS01358; ZF_RANBP2_1; 4. PROSITE; PS50199; ZF_RANBP2_2; 4. Nuclear protein; Transport; Repeat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR001876; Znf-RanBP.
Pfam; PF00641; Zf-RanBP; 4.
SMART; SM00547; ZnF_RBZ; 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          152823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                Rattus norvegicus (Rat).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15
446
687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                750
818
875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       442
657
657
721
789
846
846
1468 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                         NCBI_TaxID-10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Phosphorylation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches 144;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ZN_FING
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ZN_FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N_FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IN FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1052
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 308
음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
```

Page

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      131 GGLTVLGMLGGLYGMFIPHETVKASGEELPLVPVDPAKKQVNWDHYGNDAGGSRFVALDQ 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                76 INTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDP 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   191 INRNN----VSKLKEAWRFRTGDFTTGTGNGAEDQMTPLQ------VGNKVFLCT 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     236 PHNNIFAIDADSGKQLWKAEVNSTADAWERCRGVAYFDSTQPLVQPTLAGATPVAALAAN 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                296 TECPRRVYTNTVD-GRLIAVNADTGARCKDFGVNGTV----NLHEGLGENTKAPRFEVT 349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    250 VNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDT------ANDYFGYS 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    673
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         542 APGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF- 600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 30 GILYVRTDIGGMY-----RWDAANGRWIPLL-----DWVGWNNWGYN--GVVSIAADP 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----TIFVG-----VADPNNPVFWSRD---GGATWQAVPGAPTGFIPHKGVFDP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SAPTIAGTTIVVGSRIADNVAADMPGGVIRAYDVITGKLRWAFDPRNP-----DP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -NYVL-------KPGEIYKRSSTN-SWAAMSYDPQMNTVFLPMGSSSVDVWGGN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATW-----TRIWDWTSYPNRSLRYVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DISAEPWLT-FGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   508 SGOFYVLDRVTGKPLTKVIEQPIKVADIPGEQYSKTQPRSVEMPQIGNQTLKE-SDMWGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                567 TPFDQLMCRINFKSMRYDGLYTAP----GTDVSLSFPGSLGGMNWGSI----AFDPTHRY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -------PNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFVW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               619 MFVNDMRLGLWIQLIKQTPED -- IKIQANGGEKVNTGM -- GAVPMKGT - PYKVNKNRFMS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          161 ----WSQMTNFPDVGTYIANPIDTTGYQSD--IQGVVWVAFDKSSSSLGQASK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         445 RTAADHKY-NTSVLALD-----ATTGKEKWVYNTVHNDLWDF-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GGQIHI-----APMVKGLEE------TAVNDLISPPSGAPLISALGDLGGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           THADVTAVP-----STIFTSPVFTTGTSVDYA-----ELNPSIIVRAGSFDPSSQ-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           306;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---TDSGAT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 1; Length 809;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                       EMBL; 105770; AAC37161.1; -
InterPro; IPR001479; Bac_PQQ.
InterPro; IPR001479; Bac_PQQ.repeat.
Pfam; PF01011; Bacterial_PQQ; 7.
PROSITE; PS00363; BACTERIAL_PQQ_1; FALSE_NEG.
PROSITE; PS00364; BACTERIAL_PQQ_2; FALSE_NEG.
Oxidoreductase; PQQ; Quinate metabolism; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                           71F67CEBEA62BFCB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. 0.059;
90; Mismatches 254;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      481 DLPMQPSLVDFPMKDGTTKPAVVIG--------
                                                                                                                                                                                                                                                                                                                                             POTENTIAL. POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3.9%; Score 157;
                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             136 NNDNILYFGAPSGKGLWRS------
                                                                                                                                                                                                                                                                                                                                                                                                           MM;
                                                                                                                                                                                                                                                                                                                                                                                                         88196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 20.5
Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                         809 AA;
                                                                                                                                                                                                                                                                                                                      41
68
90
127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                               TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                  TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                       TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        500
    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

```
PTM: EXTENSIVELY O-LINKED GLYCOSYLATED ON SER AND THR RESIDUES OF THE REPEAT UNITS. HIGHEST GLYCOSYLATION APPEARS TO OCCUR ON SER RESIDUES WHICH HAVE GLY AT POSITIONS AT +2 OR -2 FROM THE GLYCOSYLATION SITE OR, WHERE GLY IS THE PENULTHMATE RESIDUE. THE PRESENCE OF PROLINE (USUALLY AT POSITION +3 OR -3) APPEARS TO ALSO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gerken T.A., Owens C.L., Pasumarthy M.; Gerken T.A., Owens C.L., Pasumarthy M.; Determination of the site-specific O-glycosylation pattern of the portion submaxillary mucin tandem repeat glycopeptide. Model proposed for the polypeptide:galnac transferase peptide binding site."; J. Biol. Chem. 272:970-9719(1997).
-!-FUNCTION: APOMOTIN IS PART OF MUCIN, THE MAJOR GLYCOPROTEIN SYNTHESIZED AND SECRETED BY MUCOUS CELLS OF THE SUBMAXILLARY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLAND. ITS HIGHLY VISCOUS AQUEOUS SOLUTIONS SERVE TO LUBRICATE THE ORAL CAVITY AND TO PROTECT IT FROM THE EXTERNAL
                                                            --- COKPPFGTMTAIDMKTRQVAWQVPLGTIQDTGPMGIKMG
                                    --QPVAAGLPSSGAV----GVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGV-S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUBUNIT: INTERMOLECULAR DISULFIDE BONDS COULD HELP MAINTAIN A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MULTIMERIC MUCIN STRUCTURE.
-!- SUBCELLULAR LOCATION: Secreted.
-!- TISSUE SPECIFICITY: SUBMAXILLARY MUCOSAE.
-!- DOMAIN: CONTAINS TANDEMLY REPEATED, IDENTICAL SEQUENCES OF 81
                                                                                                                                                                                                                                                                                                                                                                                          Euteleostomi;
Sus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eckhardt A.E., Timpte C.S., Abernethy J.L., Toumadje A., Johnson W.C. Jr., Hill R.L.; "Structural properties of porcine submaxillary gland apomucin."; J. Biol. Chem. 262:11339-11344(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Timpte C.S., Eckhardt A.E., Abernethy J.L., Hill R.L.; "Porcine submaxillary gland apouncin contains tandemly repeated, deptrical sequences of 81 residues."; J. Biol. Chem. 263:1081-1088(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eckhardt A.E., Timpte C.S., Abernethy J.L., Zhao Y., Hill R.L.; "Porcine submaxillary mucin contains a cystine-rich, carboxyl-terminal domain in addition to a highly repetitive,
                                                                                                                                                                                                                                                                                                                                                                                          Metazoa; Chordata; Craniata; Vertebrata;
Eutheria; Cetartiodactyla; Suina; Suidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYDRATE-LINKAGE SITES, AND SEQUENCE OF 45-125
                                                                                                                 069
                                                                                                                                                   OGTPMSYMSPKTGKOY - - - - VVVSAG - - - GAROSPDHG 799
                                                                                                                                                                                                                                                                                                             Last sequence update)
Last annotation update)
                                                                                                                 653 SAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCG
                                                                                                                                                                                                                                                       1150 AA
                                                                                                                                                                                                                                                                                                                                                      Apomucin (Mucin core protein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Biol. Chem. 266:9678-9686(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TISSUE-Submaxillary gland;
MEDLINE-88087170; PubMed-2826455;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TISSUE=Submaxillary gland;
MEDLINE-87280230; PubMed=3611111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TISSUE=Submaxillary.gland;
MEDLINE=91236743; Pubmed=2033060;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TISSUE=Submaxillary gland;
MEDLINE=97248516; PubMed=9092502;
                                                                                                                                                                                                                                                                                         01-OCT-1989 (Rel. 12, Created)
                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota; Metazoa; Chordata;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE OF 1-503 FROM N.A.
                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                    (Rel. 40,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       glycosylated domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 45-80.
                                                                                                                                                                                                                                                                                                                                                                    Sus scrofa (Pig).
                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=9823;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ENVIRONMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESIDUES
                                                                                                                                                                                                                                                                                                               01-DEC-1992
674 ALGIP-
                                                                                                                                                                                                                                                                                                                                16-OCT-2001
                                                                                                                                                                                                                                                   APMU_PIG
P12021;
                                                                                                                                                       169
                                        601
                                                                                                                                                                                                                                                                         Q
                                                                        g
                                                                                                               ò
```

```
AMYB_PAEPO
P21543;
 CARBOHYD
              CARBOHYD
                         CARBOHYD
                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AMYB_PAEPO
                                                                                                                                157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FFFS
                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                             음
                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                       δy
                                                                                                                                                                                                                                                                                                                                                                                                                 엄
                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             අ
                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
          SIMILARITY: CONTAINS 1 VWFC DOMAIN.
SIMILARITY: CONTAINS 1 C-FERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MUCIN TYPE)
(MUCIN TYPE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (MUCIN TYPE)
(MUCIN TYPE)
(MUCIN TYPE)
(MUCIN TYPE)
(MUCIN TYPE)
(MUCIN TYPE)
(MUCIN TYPE)
(MUCIN TYPE)
(MUCIN TYPE)
(MUCIN TYPE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE).
TYPE).
TYPE).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE)
TYPE)
TYPE)
TYPE)
TYPE)
TYPE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MUCIN TYPE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MUCIN TYPE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL)
                                                                                                                                                                                                                                                                                                                         AA TANDEM REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BY SIMILARITY.

BY SIMILARITY.

O-LINKED (GALNAC.

O-LINKED (GALNAC.)

O-LINKED (GALNAC.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GALNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GALNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GALNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GALNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GALNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GALNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GALNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GALNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GALNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GALNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GALNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GALNAC
                                                                                                                                                                                                                                                                                                                                                                                                      (INCOMPLETE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY. SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY. SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                O-LINKED
O-LINKED
O-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        O-LINKED
O-LINKED
O-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                O-LINKED
O-LINKED
O-LINKED
O-LINKED
O-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     O-LINKED
O-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          O-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                O-LINKED
                                                                                                                                                                              PIR; A40009; A40009.
InterPro; IPR000359; Cys_knot.
InterPro; IPR01007; VWFC.
Ffam; PF00007; Cys_knot; 1.
SMART; SM00041; CT; 1.
SMART; SM00214; VWC; 1.
                                                                                                                                                       EMBL; M61883; AAA30998.1; -. EMBL; M21174; AAA30990.1; -.
                                                                                                                                                                                                                                                        PROSITE; PS01185; CTCK_1; 1. PROSITE; PS01225; CTCK_2; 1. PROSITE; PS01208; VWFC; 1.
 ENHANCE GLYCOSYLATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1146
1109
                                                                                                                                                                                                                                                                                                                                                                                                                                                      1123
1139
1141
1145
                                                                                                                                                                                                                                                                                                  Repeat.
                                                                                                                                                                                                                                                                                                                         <1
<1
45
126
207
208
369
929
1062
1062
1076
1085</pre>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Glycoprotein; R
NON TER
DOMAIN
<1
REPEAT
<126
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
369
REPEAT
369
                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
DOMAIN
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DISULFID
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                         DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
```

```
31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         603
                                                                                                                                                                                                                                                                                                                                                             Query Match 3.8%; Score 153; DB 1; Length 1150;
Best Local Similarity 21.7%; Pred. No. 0.15;
Matches 147; Conservative 75; Mismatches 261; Indels 194; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKS----SSSLGQASKTIFVG 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           182 SGSTGSSSGSPGATGASIGQPETSRISVAGSSGAPAVSSGASQAAGTSGAGPGTTASSVG 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  213 VADPNNPVFWSRDGGATWQAVPGA------PTGFIPHKGVFDPVNHVLYIATSN 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  261 TGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQIS 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       294 SGAPAVSSGASQAAGTSGAGPGTTASSVGVTETARPSVAGSGTT-----GTVSGASGST 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         321 WWPDTIIFRSTDG--GATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKL 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 379 GWMDEAMAIDPFNSDRMLYGTGATLY-ATNDLTK---------WDSGGOIHIAP 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        423 MVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSP-----VFTTGTS 477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----ELNPSIIVRAGSFDPSSQPNDRHV 505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     665 SNSVPSGGS-----GTPG-----SGI----NTGGSSTQVTGIQTGTTAVGFGST 704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-MAY-1991 (Rel. 18, Created)
01-MAY-1991 (Rel. 18, Last sequence update)
01-MAY-2002 (Rel. 41, Last annotation update)
10-MAR-2002 (Rel. 41, Last annotation perupy)
10-MAR-2002 (Rel. 41, Last annotation perupy)
10-MAR-2002 (Rel. 41, Last annotation perupy)
11-MAR-2002 (Rel. 41, Last annotation perupy)
11-MAR-2002 (Rel. 41, Last annotation perupy)
12-MAR-2002 (Rel. 41, Last annotation perupy)
13-MAR-2002 (Rel. 41, Last annotation p
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                348 -----GSSSGSPGATGASI----GPETSRISVAGSSGAPAVSSGA-----SQAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        449 TSIGRSGTTHTDL---PGGTTIV-----LPGFSHSSQSSKPGSSVTTPGSPESGSETGTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         545 DPGQPVVYAVGFG-NSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                604 AAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVN-VGFGKS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              663 -APGSSYPAVF------VVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   561 GPQT-GAQTTVPGSQVSGSETGTSEAVSNPALASGSSSTGTTSGASDSQVTGSR----TG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  242 VTETARP---SVAGSGTTGTVSGASGSTGSSSGSPGATG----ASIGOPETSRISVAGS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            501 GEFSTTVISGSSHTEATTFIGGSGSPGTGSRPGTTGELSGTTIASGNATTEATTSTETRI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            506 AFSTDGGKNWFQGSEPGGVTTG-------GTVAASAD----GSRFVWAPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    616 TTGVVLGTTVAPGSSSTGATTGVLINEGTRSTSLGTT---RVASGTTYES-----GT
N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
dw. 3CB68B5D29DD7F5A CRC64;
                                                                                                                                                                                     MM.
985 985
1002 1002
1068 1068
1150 AA; 109615 M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     764 IVITG-----VTVGT 773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     708 QAITGDHANLRRVYIGT 724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     478 VDYA-----
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                              "cloning and nucleotide sequence of the gene coding for enzymatically active fragments of the Bacillus polymyxa beta-amylase.";
J. Bacteriol. 169:1564-1570(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           polymyxa beta-amylase.";
Biochemistry 30:4594-4599(1991).
-!- FUNCTION: THE PRECURSOR PROTEIN IS PROTEOLITICALLY CLEAVED TO
PRODUCE MULTIFORM BETA-AMYLASES AND A 48 KDA ALPHA-AMYLASE AFTER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CATALYTIC ACTIVITY: Hydrolysis of 1,4-alpha-glucosidic linkages
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Uozumi N., Sakurai K., Sasaki T., Takekawa S., Yamagata H., Tsukagoshi N., Udaka S.; Sasaki T., Takekawa S., Yamagata H., As single gene directs synthesis of a precursor protein with betand alpha-amylase activities in Bacillus polymyxa."; J. Bacteriol. 171:375-382(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-91215008; PubMed-1827035;
Uozumi N., Matsuda T., Tsukagoshi N., Udaka S.;
"Structural and functional roles of cysteine residues of Bacillus
                                                                                                                                                                                                                                                                                                                            MEDLINE-87231094; PubMed-2438660;
Rhodes C., Strasser J., Friedberg F.;
"Sequence of an active fragment of B. polymyxa beta amylase.";
Nucleic Acids Res. 15:3934-3934(1987).
                             MEDLINE-87165765; Pubmed-2435707;
Kawazu T., Nakanishi Y., Uozumi N., Sasaki T., Yamagata H.,
Tsukagoshi N., Udaka S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Multifunctional enzyme; Hydrolase; Glycosidase; Signal; Polysaccharide degradation; Repeat.
                                                                                                                                                                                                                                                                                                                                                                                                               DISULFIDE BOND, AND MUTAGENESIS OF CYSTEINE RESIDUES
                                                                                                                                               SEQUENCE OF 689-1196 FROM N.A., AND PARTIAL SEQUENCE
SEQUENCE OF 1-936 FROM N.A., AND PARTIAL SEQUENCE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BETA/ALPHA-AMYLASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BETA-AMYLASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PIR; A29130; A29130.
PIR; A29108; A39108.
PIR; A32251, A32251.
InterPro; IPR000461; Alpha_amylase.
InterPro; IPR001554; Glyco_hydro_14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROSITE; PS00506; BETA_AMYLASE_1; 1.
PROSITE; PS00679; BETA_AMYLASE_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF02806; alpha-amylase_C; 1.
Pfam; PF01373; Glyco_hydro_14; 1.
PRINTS; PR00750; BETAAMYLASE.
                                                                                                                                                                               MEDLINE-89123046; PubMed-2464578;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF00128; alpha-amylase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; M15817; AAA85446.1; -. EMBL; Y00150; CAA68344.1; -.
                                                                                                                                                                                                                                                                                             SEQUENCE OF 1-776 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       454
558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SECRETION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
DOMAIN
REPEAT
                                                                                                   active
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIGNAL
```

```
45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           103 KEAGLKWVPIISTHKCGGN-----VGD-----DCN-----IPLPS-----WLSSKGSADEM 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 NNWG-----YNGVVSIAAD-----PINTNKV-WAAVGMYTNSWDPNDGAILRS 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           350 SLRYVLDISAEPWLTF-----GVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGA 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                C->S: 5 FOLD DECREASE IN ACTIVITY.
C->S: 60 FOLD DECREASE IN ACTIVITY.
C->S: 60 FOLD DECREASE IN ACTIVITY.
M -> MIGL (IN REF. 3).
N -> S (IN REF. 3).
N -> D (IN REF. 3).
E -> Q (IN REF. 3).
N -> CIN REF. 3).
N -> C (IN REF. 3).
N -> S (IN REF. 3).
D -> A (IN REF. 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               :: | : | : | : | : | 343 LIQKFKDADLD--LTFTCLEMSDSGTAPNYSLPS------TLVDTVSS--IANAKGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               103 SDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            144 OF------KDESGYANSEALSPLWSGTGKQYDELYASFAENFAGYKSIIPKIY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              265 YDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            240 LDKINA-AWGTKLTS--LSQINPPTDGDGFYTNGGY-----NSAYGKDFLSWYQS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----IWDWTSYPNR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          402 TLYATNDLTKWDSGGQIHIAPMVK----GLEETAVNDLISPPSGAPLISALGDLGGF---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----THADVTAVPSTIFTSPVF------TTGTSVDYAELN--PSIIV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           445 IISKAKPDNNGGTGNKVTIYYKKGFNSPYIHYRPAGGSWTAAPGVKMQDAEISGYAKITV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QMTNFPDVGTYIANPTDTTGY-QSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             222 WSRDGGATWQAVPGAPTGFI-----PHKGVF----DPVNHVLYIATSNTGGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ------GGPSGELRYPSYYPAAGWSYPGRGKFQAYTETAKNAFRTAMNDKYGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           390 RLNGENALPTGGSGFQKIEEKITKFGYHGFTLLRINNLVN-NDGSP----TGELSGFKQY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3.8%; Score 124, ...
21.1%; Pred. No. 0.18;
tive 72; Mismatches 282; Indels 270;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A41EA6B70F257064 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 152; DB 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N -> K (IN REF.
N -> S (IN REF.
S -> N (IN REF.
MW; A41EA6B70F25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---IFRSTDG-----GATWTR--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            490 RAGSFDPSSQPNDRHVAFSTDGGKNWFQGS----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                130893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
  668
1196
126
124
198
118
126
358
                                                                                                                                                                                                       1000
1000
1154
1177
1228
3330
3330
5532
6681
6681
728
736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                               330
425
493
493
532
559
665
681
681
725
741
741
758
                                                                                                                                                                                                                      67
100
154
177
227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches 167;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            325 TI----
REPEAT
DOMAIN
DISULFID
                                                                       ACT_SITE
ACT_SITE
MUTAGEN
MUTAGEN
MUTAGEN
                                                                                                                                                                                             CONFLICT
CON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             191 LS·
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           455
  Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
```

us-09-917-376-3.rsp

```
THE CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE
CHAINS, N. TINKED AND O-LINKED OLIGOSACCHARIDES.
DISEASE: DEFECTS IN AGCI ARE THE CAUSE OF CARTILAGE MARRIX
DEFICIENCY (CMD), AN AUTOSOMAL RECESSIVE SYNDROME CHARACTERIZED BY
CELET PALATE, SHORT LIMBS, TAIL AND SNOUT. MUTATION IN STRAIN CMD
CAUSES ABSENCE OF AGGRECAN BY TRUNCATION OF THE PROTEIN (MUTATION
IN THE G1 DOMAIN).
SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY).

DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO
TERMINDS OF THE PROTEGGLYCAN, WHILE ANOTHER GLOBULAR REGION,
MAKES UP THE COOH TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS
CONSISTS OF THREE DISCLEIDE-BONDED LOOPS STRUCTURES DESIGNATED AS
THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULEPATE (KS)
AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2
                                                                                                                                                                                                                                                        PGCA_MOUSE STANDARD; PRT; 2132 AA.
661282; 064021.
10.NOV-1997 (Rel. 35, Last sequence update)
01.NOV-1997 (Rel. 35, Last annotation update)
84grecan core protein precursor (Cartilage-specific proteoglycan core protein) (CBPCP).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       the aggrecan gene.";
Nat. Genet. 7:1454-157(1994).
Nat. SPOTECGLYCAN IS A MAJOR COMPONENT OF THIS PROTEIN IS TO RESISTS TOOMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO HYALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION. MAY PLAY A REGULATORY ROLE IN THE MATRIX ASSEMBLY OF THE CARTILAGE.
-!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX (BY
                                                                                                                       558 GA----NPGDGGGTTNKVTVYYKKGFNSPYIHYRPAGGSWTAAPGVKMQDAEISGYAKITV 614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mouse cartilage matrix deficiency (cmd) caused by a 7 bp deletion in
 -----GNSWAASQGVPAN-----AQIRS 574
                                                    DRVNPKTFYAL-----SNGT-FYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKE 624
                                                                             615 DIGSASQLEAAFNDGNNNWDSNNTKNYLFSTGTSTYTP----GSNGAAGTIRTGAP--S 667
                                                                                                         GDLWLAASSGLY - - - HSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Watanabe H., Kimata K., Line S., Strong D., Gao L.-Y., Kozak C.A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   "Complete coding sequence, deduced primary structure, chromosomal localization, and structural analysis of murine aggrecan."; Genomics 22:364-371(1994).
                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-BALB/C; TISSUE-Cartilage;
MEDLINE-95104847; Pubmed-7806222;
Walcz E., Deak F., Erhardt P., Coulter S.N., Fueloep C., Horvath
Doege K.J., Glant T.T.;
537 SRFVWAPGDPG-----QPVVYAVGF---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN=129/SV;
MEDLINE=95004579; PubMed=7920633;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 211-326 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                     AGC1 OR AGC.
Mus musculus (Mouse)
                                                                                                                                                                             :[|: | | ||
723 SGAFTSMDLGT 733
                                                                                                                                                              681 TGAYRSDDCGT 691
                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Yamada Y.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AND G3
                                                                                                                                                                                                                                             PGCA_MOUSE
                                                    575
                                                                                                         625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         +
                         g
                                                                          q
                                                                                                       δ
                                                                                                                                 g
                                                                                                                                                              à
                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                      à
                                                    ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
-i- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
-i- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
-i- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC. . ) (POTENTIAL).
(GLCNAC. . ) (POTENTIAL).
(GLCNAC. . ) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PROSITE; PS00290; IG_MHC; 1.
PROSITE; PS00241; LINK; 4.
PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
Glycoprotein; Cartilage; Proteoglycan; Lectin; Signal; Sushi; Stepat; Immunoglobulin domain.
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AGGRECAN CORE PROTEIN.
IG-LIKE V-TYPE DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LINK 4.
C-TYPE LECTIN.
SUSHI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY.
SIMILARITY.
SIMILARITY.
SIMILARITY.
SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIM
                                                                                                                                                                                                                                                                                                                                                                  EMBL; L07049; AAC37670.1; --
EMBL; S73721; AAB32160.1; --
EMBL; S73721; AAB32160.1; --
HSSP; P98066; 1756.
MGD; MGI:99602; AGC.
InterPro; IPR003006; Ig_MHC.
InterPro; IPR003324; SGXXSG.
InterPro; IPR0004356; Ig_V.
InterPro; IPR0004356; Ig_V.
InterPro; IPR0004356; Link.
InterPro; IPR001304; Link.
InterPro; IPR00104; Lectin_C.
Pfam; PF00099; Lectin_C.
Pfam; PF00099; SGXXSG; 62.
Pfam; PF00099; XIINK; 4.
FP000m; PR000918; Link; 4.
SMART; SM00031; CCP; 1.
SMART; SM00046; IGV; 1.
SMART; SM000466; IGV; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LINK 1.
LINK 2.
LINK 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CS-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  G1-A.
G1-B.
G1-B'
G2-B.
G2-B'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2132
140
247
247
349
581
683
2044
2106
140
247
349
580
682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1231
1917
2132
133
133
246
220
220
348
318
580
584
682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  233
233
233
233
333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
```

σ

```
CHAIN
                                                                                                                                                                    39;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1532
                                                                                                                                                                                                                                                                                                                                                                                                                      1267 D-------FSGLPSGF-PTVSLVD--STLVEVITATTSSELEGRGTIGIS 1306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1591 EAGEGPSGILEFSGAHSGTPDISGELSGSLDLSTLOSGOMETSTETPSSPY--FSGDFSS 1648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1649 TTDVSGESIAATTGSGESSGLPEVTLNTSELVEGVTEPTVSQELGHGPSMTYISRLSEAS 1708
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1709 GDASASGDLGGAVTNFPGSGVEASVPEASSDLSAYPEAGVGVSAAPEASSKLSEFPDLHG 1768
                                                                                                                                                                                                                                   GYVSGIPSGGDGTET ---- SASGVEDVSGLPSGGEGLETSASGVEDLGPSTRDSLETSAS 1160
                                                                                                                                                                                                                                                                                         1216 VGSASGALDFGKLPPGTLGSGQTPEVNGFPSGFSG-----EYSGADIGSGPSS---GLP 1266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1307 GSGEV-----SGL-----PLGELDSSAD---ISGL-----PSGTELSGQASGSPDS-- 1344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SGETSG------EFTSGI 1371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       468
                                                                                                                                                                                                                                                                                                                                                                                                      267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---EET 430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       516
                                                                                                                                                                    Gaps
                                                                                                                                                                                                  64 GYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGA------TWQITPL 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1769 ITSAFHETDLEMTTPSTEVNSNPWTF--QEGTREGSAAPEVSGESSTTSDIDTGTSG 1823
                                                                                                                                                                                                                                                                                                                                  -------VGTYIANPT-DTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   328 FRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1489 GSVELSGFPSGETELSGTSGTVDVSEQSSGAIDSSGLTSPTPE-----F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AVGFGNSW----AASQGVP-----ANAQIRSDRVNPKTFYALSNGTFYR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --STNGGSSWSAIT-----GVSSAV------NVGFGKSAPGSSYPAVFVVGTIGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   268 SSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----SPVFTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWF
                                                                                                                                                                                                                                                                  115 PFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPD----
                                                                                                                                                                                                                                                                                                                                                                                                   DPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    431 AVNDLISPPSGAPL-----ISALGDLGGFTHADVTAVPSTIFT----
 N-LINKED (GLCNAC. . .) (POTENTIAL).
CELL ATTACHMENT SITE (POTENTIAL).
MM: OB2BCDFCGCBDA163 CRC64;
                                                                                                                                  3.7%; Score 149; DB 1; Length 2132;
19.5%; Pred. No. 0.53;
tive 96; Mismatches 286; Indels 292;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              680 VTGAYRSDDCGTT------WVLINDDQHQYGNWGQAITGDHANLRRVYIGTNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   388 DPFNSDRMLYGTGATLYATNDLTKWD--SGGQIHIAPMVKGL------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STD-GGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYH-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QGSEPG-----GVTTGGTVAASA-DGSRFVWAPGDP-----
387
611
667
1675
1173
AA; 222008 M
                                                                                                                                                                   Conservative
                                                                                                                                                     Best Local Similarity
 387
611
667
1675
1171
                                                                                                                                                                   Matches 163;
 CARBOHYD
CARBOHYD
                                  CARBOHYD
CARBOHYD
                                                                                   SEQUENCE
                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ISOA_PSEAY
                                                                                                                                                                                                                                                                                                                                  170
                                                                                                                                                                                                                                                                                                                                                                                                   215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          593
                                                                                                                                                                                                                                   1105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            638
                                                                     SITE
   FFFF
                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ΩD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
```

776 AA

STANDARD;

ISOA_PSEAY P10342;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                branch
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             J. Mol. Biol. 281:885-897(1998).

-!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1,6)-D-glucosidic branch linkages in glycogen, amylopectin and their beta-limits dextrins. IntRages in MALTOSE.

-!- INDUCTION: BY MALTOSE.

-!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY.
                                                                                                                                                                                                                                                                           Chen J.H., Chen Z.Y., Chow T.Y., Chen J.C., Tan S.T., Hsu W.H.; "Nucleotide sequence and expression of the isoamylase gene from an isoamylase-hyperproducing mutant, Pseudomonas amyloderamosa JD210."; Biochim. Biophys. Acta 1087:309-315(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                           X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
MEDLINE-993887985; PubMed=9719642,
Katsuya Y., Mezaki Y., Katsuya Y.;
"Three-dimensional structure of Pseudomonas isoamylase at 2.2-A
                                                                                                                                                                   M., Noumi T., Futai M.;
f the isoamylase gene from
                                                                                                                                                                                                                                                                                                                                                                                          Amemura A., Fujita M., Futai M.;
"Transcription of the isoamylase gene (iam) in Pseudomonas
amyloderamosa SB-15.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A -> G (IN REF. 1).
F -> C (IN REF. 1).
GAS -> AH (IN REF. 1).
           15-DEC-1998 (Rel. 37, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
Isoamylase precursor (EC 3.2.1.68).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hydrolase; Glycosidase; Signal; 3D-structure. SIGNAL 1 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY.
SIMILARITY.
SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          I SOAMY LASE,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-88243808; PubMed=3379068;
Amemura A., Chakraborty R., Fujita M
"Cloning and nucleotide sequence of
Perdodomonas amyloderamosa SB-15";
J. Biol. Chem. 263:9271-9275(1988).
                                                                                                                                                                                                                                                                                                                                                               STRAIN-SB-15;
MEDLINE-89327147; PubMed-2753857;
                                                                                                                                                                                                                                                                                                                                                                                                                                   Bacteriol. 171:4320-4325(1989)
                                                                                                                                                                                                                                                                 MEDLINE-91064385; PubMed-2248978;
  Created)
                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 744-776 FROM N.A.
                                                                      Pseudomonas amyloderamosa
                                                                                Bacteria; Proteobacteria.
NCBI_TaxID=32043;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       401
481
536
422
616
766
01-MAR-1989 (Rel. 10,
15-DEC-1998 (Rel. 37,
15-DEC-1998 (Rel. 37,
                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       401
481
536
536
410
546
738
126
126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       resolution.";
                                                                                                                                          STRAIN-SB-15;
                                                                                                                                                                                                                                                     STRAIN-JD210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
ACT_SITE
ACT_SITE
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DISULFID
```

```
42;
                                                                                                                                                                                                                                                                                                                                         AQEVSQDP-----LNPSNQ------NGNVFASGASYR-----TIDSGIY--AP 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 503 YVTQDANDFSGS-----SNLFQSSGRSPWNSINFIDVHDGMTLKDVYSCNGANNSQAWP 556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              599
                                                                                                                                                                                                                                                                                                                                                                                  201
                                                                                                                                                                                                                                                                                                                                                                                                                                                  LGQASKTIFVGVAD-----PN---NPVFW-----SRDGGATWQAVPGAPTG-F 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    302 QAMVQAFHNAGIKVYMDV------VYNHTAEGGTWTSSDPTTATIYSWRGLDN 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DWTS-YPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAID----PFNSDRML 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLR 718
                                       SGLDLFAEPWAIGGNSYQLGGFPQGWSEWNGLFRDS -> TVWICLRNLGPSAATRISWVDSRRVVRVEWSVPRQ (IN
                                                                                                                                                                                                                                            27 GAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVG 86
L -> V (IN REF. 1).

GAYT -> AVH (IN REF. 1).

SGLDLFAEPWAIGGNSYQLGGFPQGWSEWNGLFRDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ALRPSS---------WYSGSQLTWYQPSGAVADSNYWNNTSNYAIAYAIN--
                                                                                                                                                                                                                                                                              GITGAVY ----YGYRAWGPNWPYASNWGKGSQAGFVSDVDANGDRFNPNKLLLDPY
                                                                                                                                                                                                                                                                                                              87 MYTNSWDPNDCAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAP
                                                                                                                                                                                                                                                                                                                                                                              SGKGLWRSTDSGATWSQMTNFPDV - GTYIANPTD-TTGYQSDIQGVVWVAFDKSS--SS
                                                                                                                                                                                                                                                                                                                                                                                                                183 KGVVLVPSTQSTGTKPTRAQKDDVIYEVHVRGFTEQDTSIPAQYRGTYYGAGLKASYLAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  243 LG-VTAVEFLPVQETQNDANDVVPNSDANQNYWGYMTENYFSPDRRYAYNKAAGGPTAEF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----DTANDYF-GYSGLTIDRQHPNTI---MVATQISWWPDTIIFRSTDGGATWTRIW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      401 DLASVLGNSCLNGAYTASA-----PNCPNGGYNFDAADSNVAINRILREFTVRPAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        397 YGTGATLYATNDLTKWDSGGQIH-IAPMVKGLEE------TAVNDLISPPSGAPLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGSGLDLFA----EPWAIGGNSYQLGGFPQGWSEWNGLFRDSLRQAQNEL----GSMTI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SALGDLGGFTHADVTAVPST1FTSPVFTTGTSVDYAELNPS11VR-----AGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            557 YGPS------DGGTSTNYSWDQGMSAG---TGAAVDQRRAARTGMAFEML--SA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               547 GQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTF-YALSNGTFYRSTDGGVTFQPVAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GTPLMQG----GDEYLRTLQCNNNAYNLDSSANWLTYSWTTDQSNFYTFAQRLIAFRKAHP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GL-PSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNG----GSSWSAIT--GVSSAVNVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -GPSLGDSNSIYVAYNGWSSSVTFTLPAPPSGTQWYRVTDT----CDWNDGASTFVAPGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    241 IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPST------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FDPSSQPNDRHVAFSTDGGK - - - - NWFQGSEPGGVTTGGTV - - - AASADGSRFVWAPGDP
                                                                                                                                                                                                            213;
                                                                                                                                                                       Length 776;
                                                                                                         Ξ.
                                                                                   WP -> S (IN REF. 1).
AFRKAHPA -> RSARHIP (IN REF. F738BF8040246169 CRC64;
                                                                                                                                                                         ; Score 148.5; DB 1; Length 7; Pred. No. 0.17; 90; Mismatches 330; Indels
                                                                       REF. 1).
                                                                                                                       83626 MW;
                                                                                                                                                                         3.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RVYIGTNGRGIVYGDIG 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    753 ETLIG -- GAGTTYGQCG 767
                                                                                                                                                                                                           Matches 164; Conservative
   386
416
489
                                                                                     556
657
                                                                                     555
650
776 AA;
                                                                                                                                                                                         Similarity
   386
413
454
                                                                                   CONFLICT
CONFLICT
SEQUENCE
 CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                         Query Match
                                                                                                                                                                                            Local
                                                                                                                                                                                                                                                                                                                                                                              147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   629
   FTFFFF
                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Dp
                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                              οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

RESULT 9 ISOA_PSESP

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-s1b.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     42;
                                                                                                                                                                                                                                                                                        of Pseudomonas sp.";
J. Gen. Microbiol. 135:37-45(1989).
-- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1,6)-D-glucosidic branch
--- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1,6)-D-glucosidic branch
--- InhAgges in glycogen, amylopectin and their beta-limits dextrins.
--- INDUCTION: BY MALTOSE.
--- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
                                                                                                                                                                                                                                                                           'Cloning and nucleotide sequence of the isoamylase gene from a strain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 148 AQEVSQDP-----TUPSNQ------NGNVFASGASYR-----TTDSGIY--AP 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LGQASKTIFVGVAD-----PN---PN---NPVFW-----SRDGGATWQAVPGAPTG-F 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         27 GAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVG 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           96 GITGAVY-----YGYRAWGPNWPYASNWGKGSQAGFVSDVDANGDRFNPNKLLLDPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                183 KGVVLVPSTQSTGTKPTRAQKDDVIYEVHVRGFTEQDTSIPAQYRGTYYGAGLKASYLAS
                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY. ISOAMYLASE SUBFAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               87 MYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LG-VTAVEFLPVQETQNDANDVVPNSDANQNYWGYMTENYFSPDRRYAYNKAAGGPTAEF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   241 IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPST-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     147 SGKGLWRSTDSGATWSQMTNFPDV--GTYIANPTD-TTGYQSDIQGVVWVAFDKSS--SS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels 213;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 776;
                                                                                                                                                                                                                                     Pognoni A., Carrera P., Galli G., Lucchese G., Camerini B.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A5E4C02EF026A3A4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3.7%; Score 148.5; DB 1;
20.6%; Pred. No. 0.17;
ive 90; Mismatches 330;
                                  01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
Isoamylase precursor (EC 3.2.1.68).
776 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIMILARITY.
SIMILARITY.
SIMILARITY.
SIMILARITY.
SIMILARITY.
SIMILARITY.
                                                                                                                                                                                                   SEQUENCE FROM N.A., AND SEQUENCE OF 27-44. MEDLINE=89381677; PubMed=2778432;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ISOAMYLASE.
BY SIMILARIT
BY SIMILARIT
BY SIMILARIT
BY SIMILARIT
BY SIMILARIT
BY SIMILARIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR000461; Alpha_amylase.
InterPro; IPR004193; isoamylase_N.
Pfam; PF00128; alpha-amylase; I.
Pfam; PF02922; isoamylase, I.
Hydrolase; Glycosidase; Signal.
PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MW;
                                                                                                                           Pseudomonas sp. (strain SMP1).
Bacteria, Proteobacteria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; M25247; AAA25855.1; -. EEMBL; A709999.1; -. PIR; A37035; A37035. HSSP; P10342; 18F2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 20.6%
Matches 164; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             83656
STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    481
536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           776 AA;
                                                                                                                                                               NCBI_TaxID=306;
 SOA_PSESP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DISULFID
DISULFID
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           202
ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
```

```
PERMINELLY SSCIAL2 / ATCC 700720;
MEDLINE-21534948; PubMed-11677609;
MCClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P., Courtbey L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D., Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M., "Complete genome sequence of Salmonella enterica serovar Typhimurium
                                                                                                                                                                                      445
                                                                                                                                                                                                                         502
                                                                                                                                                                                                                                                             493
                                                                                                                                                                                                                                                                                                                                                                                                           605
                                                                                                                                                                                                                                                                                                                                                                                                                                                959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           718
--VYNHTAEGGTWTSSDPTTATIYSWRGLDN 348
                                                                                                                                                DLASVLGNSCLNGAYTASA------PNCPNGGYNFDAADSNVAINRILREFTVRPAA 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              698 -GPSLGDSNSIYVAYNGWSSSVTFTLPAPPSGTQWYRVTDT----CDWNDGASTFVAPGS 752
                                                                                                                                                                                                                                                                                    503 YVTQDANDFSGS-----SNLFQSSGRSPWNSINFIDVHDGMTLKDVYSCNGANNSQAMP
                                                                                                                                                                                                                                                                                                                                                                                                                                              GTPLMQG---GDEYLRTLQCNNNAYNLDSSANWLTYSWTTDQSNFYTFAQRLIAFRKAHP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLR
                                    -----DTANDYF-GYSGLTIDRQHPNTI----MVATQISWWPDTIIFRSTDGGATWTRIW
                                                                       349 TTYYELTSGNOYFYDNTGIGANFNTYNTVAONLIVDSLAYWANTM---GVDGFR-----F
                                                                                                                                                                                                                                                           SALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVR------AGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GL-PSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNG----GSSWSAIT--GVSSAVNVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       657 ALRPSS-------WYSGSQLTWYQPSGAVADSNYWNNTSNYAIAYAIN--
                                                                                                           DWTS-YPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAID----PFNSDRML
                                                                                                                                                                                                                       452 GGSGLDLFA----EPWAIGGNSYQLGGFPQGWSEWNGLFRDSLRQAQNEL----GSMTI
                                                                                                                                                                                  397 YGTGATLYATNDLTKWDSGGQIH-IAPMVKGLEE-----TAVNDLISPPSGAPLI
                                                                                                                                                                                                                                                                                                                                  FDPSSQPNDRHVAFSTDGGK----NWFQGSEPGGVTTGGTV---AASADGSRFVWAPGDP
                                                                                                                                                                                                                                                                                                                                                                     YGPS-------DGGTSTNYSWDQGMSAG---TGAAVDQRRAARTGMAFEML--SA
                                                                                                                                                                                                                                                                                                                                                                                                           GQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTF-YALSNGTFYRSTDGGVTFQPVAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BIGA OR STM3478.
Salmonella typhimurium.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           P25927; P25928; Q9XCQ3; O1-MAY-1992 (Rel. 22, Created) 01-MAY-2002 (Rel. 41, Last sequence update) 01-MAR-2002 (Rel. 41, Last annotation update) Putative surface-exposed virulence protein bigA precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Stollijkovic'., Valentine P., Heffron F.;
"Salmonella typhimurium rhs homolog.";
Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1953 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 1-765 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  719 RVYIGTNGRGIVYGDIG 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     753 ETLIG--GAGTTYGQCG 767
 302 QAMVQAFHNAGIKVYMDV --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Nature 413:852-856(2001)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   [2]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=602;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-14028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Salmonella.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BIGA_SALTY
                                                                                                             342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         629
                                                                                                                                                                                                                                                           446
                                                                                                                                                                                                                                                                                                                                                                     557
                                                                                                                                                                                                                                                                                                                                                                                                                                            009
                                                                                                                                                401
                                                                                                                                                                                                                                                                                                                                  494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  909
                                                                                                                                                                                                                                                                                                                                                                                                         547
                                                                       g
                                                                                                             ò
                                                                                                                                              g
                                                                                                                                                                                  ò
                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                         Qγ
                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     유
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation the European Bioinformatics Institute. There are no restrictions on its by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 34;
                                                                  J. Bacteriol. 173:325-333(1991).
-!- CAUTION: Ref.3 sequence differs from that shown due to frameshifts in positions 414 and 732.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---YIVGTNA----DGSSG 1531
                       Wu J.Y., Siegel L.M., Kredich N.M.; "High-level expression of Escherichia coli NADPH-sulfite reductase: requirement for a cloned cysG plasmid to overcome limiting siroheme
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        95 NDGAILRSSDQGATWQITPLPFKLGGNMPG----RGMGERLAVDPNNDNILYFGAPSGKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       44 WDAANG-----RWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVGMYTNSWD---P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         151 LWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        211 VGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1398 NSCIINVGTEQGK------EDGTNGTGLIGIKGNGNATTINNTADGVINVYADDSYA
                                                                                                                                                                                                                                                                                                                             PUTATIVE SURFACE-EXPOSED VIRULENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 1953;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LE -> ITLQ (IN REF. 1).
-> T (IN REF. 1).
611B3F1C954D91AE CRC64;
                                                                                                                                                                                                                                                                                                                                                         15 X 11 AA TANDEM REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 274;
                                                                                                                                                                                                                                                                                                                                                                     (INCOMPLETE).
(INCOMPLETE).
(INCOMPLETE).
                                                                                                                                                                                                                                                                                                         Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3.7%; Score 148.5; Dl
20.5%; Pred. No. 0.51;
tive 83; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (INCOMPLETE)
                                                                                                                                                                                                                                                                                                                                              PROTEIN BIGA.
                                                                                                                                                                                                                                                                                                                       POTENTIAL.
                                                                                                                                                                                                                                            EMBL; M64606; AAA27042.1; ALT_FRAME
EMBL; M64606; AAA27043.1; ALT_FRAME
           MEDLINE-91100301; PubMed-1987123;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15
                                                                                                                                                                                                                    EMBL; AF133696; AAD39458.1; -. EMBL; AE008859; AAL22340.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    200150
                                                                                                                                                                                                                                                                                         Stygene; SG10437; bigA.
Virulence; Repeat; Signal;
SIGNAL 1 27
CHAIN 28 1953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 3.7
Best Local Similarity 20.5
Matches 158; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                    144
1155
1166
1177
1188
1199
1199
1221
221
232
243
                                                                                                                                                                                                                                                                     PIR; C39200; C39200.
PIR; D39200; D39200.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1698
1795
1836
1953
                                                                                                                                                                                                                                                                                                                                                         101
101
104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                       REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

```
1614
                                                                                                                                                                                                                                                                1669 MLADAAPQIKDGLAFNVV-----AKGDP----RAELGNDTQYDMLALRQTLDLTA--SQ 1716
                                                                                                                                                                                                                                                                                                                                1717 NLTLEY------GIARLDG 1729
                                                                                                                                                                                                                                                                                                                                                                                                    1787 IADSDMRQQYLEFRSEGAKTFTMMGDALKVTPYAGVKFRHTMEDGYKERSAGDFNLSMNS 1846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   675
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PGCA_CHICK STANDARD; PRT; 2109 AA.
P07895; Q90991; Q90820; Q91047; Q90810;
01-A06-1988 (Rel. 08, Created)
01-A07-1997 (Rel. 35, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Aggrecan core protein precursor (Cartilage-specific proteoglycan core
DVWKFSV-------TSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQ 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-WHITE LEGHORN; TISSUE-Embryo; MEDLINE-94043149; PubMed-8226878; Li H., Schwartz N.B., Vertel B.M.; "cDNA cloning of chick cartilage chondroitin sulfate (aggrecan) core protein and identification of a stop codon in the aggrecan gene associated with the chondrodystrophy, nanomelia."; J. Biol. Chem. 268:23504-23511(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     II.
                                                                          319 ISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLJFGVQPNPPVPSPKL
                                                                                                               1581 VVW--NAQGSQDADGNV-----DVTMTKNAYADVATDSSVS------
                                                                                                                                                  379 GWMDEAMAIDPFNSDRMLY----GTGATLYATNDLTKWDSGGO-----
                                                                                                                                                                                     1615 --- DVAQALDAGYTNNELYTSLNVGTTAEL---NSALKQVSGAQATTVFREARVLSNRFT
                                                                                                                                                                                                                            --IHIAPMVK-GLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTT
                                                                                                                                                                                                                                                                                                     GTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASA
                                                                                                                                                                                                                                                                                                                                                                             DGSRFVWAPGDPGQPVVYAVGFG--NSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYR
                                                                                                                                                                                                                                                                                                                                                                                                                                                   593 STD------GOVTFQPVAAGLPSSGAVGVMF-HAVPG--KE---GDLWLAASS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----GLYHSTNGGSSWSAITGVSSAVNVGFGKS-----APGSSYPAVFVVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Krueger R.C. Jr., Fields T.A., Mensch J.R. Jr., Schwartz N.B.; "Chick cartilage chondroitin sulfate proteoglycan core protein. Nucleotide sequence of CDNA clone and localization of the S103L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1906 KGGGVNGLATIGVKYSSND----TALHLDAYQWKEDGISDKGFMLNVKKTF 1952
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      676 TIGGVTG-----AYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVY 721
                                     1532 TLKANNLVIGDNVKVDTGFTSGTADTTVVVDNAFTGSNIQGADNIT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Biol. Chem. 265:12088-12097(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TISSUE=Embryo;
MEDLINE=90307744; PubMed=1694853;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 1042-1559 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                protein) (CSPCP).
AGC1.
Gallus gallus (Chicken).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=9031;
271
                                                                                                                                                                                                                                                                                                     475
                                                                                                                                                                                                                                                                                                                                                                             535
                                                                                                                                                                                                                            418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PGCA_CHICK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
                                     g
                                                                                                                                                                                     g
                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                               g
                                                                                                                                                δ
                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

SEQUENCE OF 1-1855 AND 1893-2109 FROM N.A.

TISSUE-Cartilage;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -!- FTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE CHAINS, N-LINKED AND O-LINKED OLIGOSACCHARIDES.
-!- DISEASE: DEFECTS IN AGC1 ARE THE CAUSE OF NANOMELIA, A LETHAL CONNECTIVE TISSUE DISORDER AFFECTING CARTILAGE DEVELOPMENT (CHONDRODYSTROPHY) CHARACTERIZED BY SHORTENED AND MALFORMED LIMBS. AGGRECAN IS TRUNCATED AT ITS C-TEMINAL IN THE CS-2 BINDING DOMAIN AND IS NOT ANYMORE SECRETED FROM THE CHONDROCYTES.
-!- SIMILARITY: CONTAINS 1 IMMUNGLOBULIN-LIKE V-TYPE DOMAIN.
-!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
-!- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
-!- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
-!- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIMILARITY).

-1- DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO TERMINUS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3, MAKES UP THE COOH TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS CONSISTS OF THREE DISULEIDE-BONDED LOOP STRUCTURES DESIGNATED AS THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS) AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-89008500, Pubmed-3170613;
Tanaka T., Har-El R., Tanzer M.L.;
"Partial structure of the gene for chicken cartilage proteoglycan
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO YALLOROUTC ACID VIA AN AMINO-TERNINAL GLOBULAR REGION. MAY REGULATORY ROLE IN THE MATRIX ASSEMBLY OF THE CARTILAGE. SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX (BY
           Chandrasekaran L., Tanzer M.L.;
"Molecular cloning of chicken aggrecan. Structural analyses.";
Biochem. J. 288:903-910(1992).
                                                                                                                                                                                              STRAIN=WHITE LEGHORN; TISSUE-Chondrocytes;
MEDLINE=951128519; PubMed-782775.
Primorac D., Stover M.L., Clark S.H., Rowe D.W.;
"Molecular basis of nanomelia, a heritable chondrodystrophy of
                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 1693-1855 AND 1893-2109 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                  J. Biol. Chem. 263:15831-15835(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              M88101; -; NOT_ANNOTATED_CDS.
S74657; AAC60751.1; -.
S74656; AAC60751.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-86259736; PubMed-3460082;
                                                                                                    MEDLINE-94107258; PubMed-8280087;
                                                                                                                                                                                                                                                                                                          [6]
SEQUENCE OF 1894-2109 FROM N.A.
                                                                                                                                                                                                                                                                                     Matrix Biol. 14:297-305(1994).
                                                                                                                       Chandrasekaran L., Tanzer M.L.
Biochem. J. 296:885-887(1993)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; L21913; AAB19128.1; -.
                                                                                                                                                           [5]
SEQUENCE OF 1492-1610 FROM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAA48719.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   M38187; AAA48731.1;
                                                                                                                                                                                                                                                                                                                                                                                                 protein.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    S74656; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AND G3
                                                                                                                                                                                                                                                                      chicken
                                                                                       ERRATUM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                   core
```

```
977 VSGEEELSGFPSA----TSGEOSGVP 1013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1014 ELSGEHSGVP----GLSGEAFEVPELSGEH----SGVT-----ELSGEHSGLPELSGEPF 1060
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  625 GDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGT--IGGVTG 682
                                                                                                                                                                                                                                                                                                                                                                                                                                                318 ----QISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPP 372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          373 ------VPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSG-----GQIH 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    420 IAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFT----HADVTAVPSTIFTSPVFTT 474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                475 GTSVDYAELNPSIIVR-----AGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGG 523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   578 ------NPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKE 624
                                                                                                                                                                                                                                                                          94 PNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNIL--YFGAPSGKGL 151
                                                                                                                                                                                                                                                                                                                                                             212 GVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGD 271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----GETAVTFPYVFGPSGEGSASGDLSGGAS 1357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         683 AYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDIGGAPS 739
                                                                                                                                                                                                                                                                                     -------SRIPEVSGMPSG------AESSGLHSGFSGE
                                                                                                                                                                                                                                                                                                                   152 WRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFV
                                                                                                                                                                                                                                                                                                                                        -----SGLPSAEEEVTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          524 VT-----TGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRV-
  (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                             1. 3.6%; Score 146; DB 1; Length 2109; Similarity 20.5%; Pred. No. 0.78; Onservative 72; Mismatches 250; Indels 248;
                                                                                                                                                                                                                                                                                                                                                                                                      272 VWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVAT-----
                                                                                                                                                      REF.
N-LINKED (GLCNAC. . .) (POTENT N-LINKED (GLN REF. 3).

A -> D (IN REF. 3).
A -> P (IN REF. 3).
A -> P (IN REF. 2).
C -> D (IN REF. 5).
I -> T (IN REF. 5).
I -> V (IN REF. 5).
I -> V (IN REF. 5).
I -> V (IN REF. 5).
I -> S (IN REF. 3).
S -> G (IN REF. 6).
                                                                                                                                                                                                          223492 MW;
                                                                                                                                                                                                                                                    147; Conservative
  644
700
700
801
1892
362
362
1000
1000
1029
1029
11594
11594
11594
11594
11594
11594
11594
11610
11610
11610
11610
11610
11610
                                                                                                                                                                                                            AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1322 PTVSQELG---
   CARBOHYD
VARSPLIC
CONFLICT
                                                                                                                                                     CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                           Local
                                                                                                                                                                                                                                                                                                                                                                                  .VS 668
                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                        875
                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ω
                                                                    (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                             AGGRECAN CORE PROTEIN.
IG-LIKE V-TYPE DOMAIN.
LINK 1.
LINK 2.
LINK 3.
LINK 4.
LINK 4.
LINK 4.
LINK 4.
C-TYPE LECTIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BY SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                               SUSHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               G1-A.
G1-B.
G1-B'
G2-B.
G2-B.
                   HSSP, P08709; IBF9.
InterPro; IPR000152; Asx hydroxyl.
InterPro; IPR000561; BGF-like.
InterPro; IPR000742; BGF_2.
InterPro; IPR011881; BGF_Ca.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            685
11870
11870
11870
11890
2017
2007
2067
2080
2080
122
330
3388
                                                                                                                                                                                                                                                                                                                                                                                                               1742
1892
2019
2081
137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DISULFID
                                                                                                                                                                                                                                                                                                                                                           DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
```

916

32;

Gaps

2

```
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=1148;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Y48L_SYNY3
P73069;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SLR2034.
                                                   862
                                                                                                                               461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Y48L_SYNY3
                                                                                                                                                                                                                                                                                     583
                                                                                                                                                                                                                                                                                                              546
 à
                         g
                                                   ò
                                                                          g
                                                                                                     ò
                                                                                                                             a
                                                                                                                                                       ò
                                                                                                                                                                               g
                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       50;
                                                                                                                                                                                                                 Lee Y.-A., LO Y.-C., Yu P.-P.;

"A gene involved in quinate metabolism is specific to one DNA homology group of Xathomonas campestris.";

J. Appl. Microbiol. 87:649-658(199).

- CATALYTIC ACTIVITY: Quinate + pyrroloquinoline-quinone = 5-dehydroquinate + reduced pyrroloquinoline-quinone.

- COFCTOR: PACHWAY: QUINIC ACID CATABOLIC PATHWAY; FIRST STEP. THIS PATHWAY ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO PROTOCATECHUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       151 WSAYGNTDGGSRFAALDQINRSNGRPAAG---SPGPTTPGEIANSDGNGAEDQLTPLQ-- 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 118 LGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTY---- 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IFVGVA-DP----NNPVFWSRD--GGATWQAVPGAPT---GFIPHKGVFDPVNHVLYIA 257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              60 WNNWGY--NGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFK 117
                                                                                                                                                                                                                                                                                                                                                               SUBCELLULAR LOCATION: Integral membrane protein (Potential). SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----VGEKVFLCTPHNNLIALDASTGKQLWR-REINATSSVWQRCRGLGYFDADA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----IANPTDTTG-----YQSDIQG-VVWV----AFDKSSSLGQASKT
                                                                                                                 Xanthomonas campestris (pv. juglandis).
Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 790;
                             16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Probable quinate dehydrogenase [Pyrroloquinoline-quinone]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PROSITE; PS00363; BACTERIAL_PQQ_1; FALSE_NEG.
PROSITE; PS00364; BACTERIAL_PQQ_2; FALSE_NEG.
Oxidoreductase; PQQ; Quinate metabolism; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    B75F29B52A49FE6F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   87; Mismatches 239;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3.6%; Score 145.5; Di
21.6%; Pred. No. 0.26;
            790 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; AF109471; AAD38453.1; -.
InterPro; IPR001479; Bac_PQQ.
InterPro; IPR002372; Bac_PQQ_repeat.
Pfam; PF01011; Bacterial_PQQ; 6.
                                                                                                                                                                                                         MEDLINE-20063481; PubMed=10594704;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    82896 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      21.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 22
48
77
106
790 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                      NCBI_TaxID=44291;
                                                                                                                                                                                                                                                                                                                                                   PATHWAY
             QUIA_XANCJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches 166;
                                                                                                                                                                                             STRAIN-C5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local
QUIA_XANCJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     209
                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its was by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb.ch/.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y., Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T., Shimpo S., Takeuchi T., Wataki A., Nakazaki N., Naruo K., Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
316 AGLGAAPDPFYQLTSPPLVAGTTVVGGRT-RADDNVQTDMPGGVVRGSMWSPVRS---A 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions."; DNA Res. 3:109-136(1996).
                                                                                  258 TSNTGGPYD-----GSS-----TDVW-KFSVTSGTWTRISPV--PSTD-----TANDY
                                                                                                                                                                 371 GLDPGNPHDRQAPAAGSSYVRSTPNVWAPMSYDAAMNTVFLPLGGPSTDLYGAERTALDH
                                                                                                                                                                                                                                                    -FGYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATW-----TRIMDWTSYPNRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      410 TKWDSGGQIHIAPMVKGLEETAVNDLISPPSGA------PLISALGDLG--GFTH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             627 EAVNTGMGAVPLKGTPYAVN-KNRFLSALGIPCQAPPYGTLSAIDLKTRSIAWQVPVGTV
                                                                                                                                                                                                                                                                                                                                                                                                                     351 LRYVLDISAEPWL-TFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         486 TK---AGQIYVLDRATGKPLTEVREV--PVKGSDIAHEQYAPTQPLSVGMPQIGTKHLTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            457 ADV---TAVPST-----IFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          541 SDMWGATAMDQMLCRIAFKQMRYEGLYTAP----GTDVSL-----SF-PGSL-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----PGQPVVYAVGFGNSWAASQGVPANAQIRS--DRVNPKT---FYALSNGTF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                591 YRSTDGGVTFQ-PVAAGLPSSGAV----GVMFHAVPGKEGDLWLAASSGLYHSTNGGSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         501 NDRHVAFSTDGGKNWFQGSEPGGVTTGGT--VAASADGSRFVWA---PGD------
                                                                                                                                                                                                                                                                                                                                      ----ATTGAEKWVYQTVHNDLWDF----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----DLPMQPSLIDFPNQDGSHTPAVVIG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Synechocystis sp. (strain PCC 6803).
Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 645 WSAITGVSSAVN-VGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15-JUL-1999 (Rel. 38, Created)
15-JUL-1999 (Rel. 38, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
YCf48-like protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ā
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             342
                                                                                                                                                                                                                                                                                                                                      431 RYGASVLALD------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
```

```
- :: :: - :: - :: - ::
                                                                                                                                                                                                                                                                                                                                                                                                                                     747
542
639
208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               747 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                     40
448
546
208
349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Simines 141;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
ACT_SITE
ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                             18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               There are no restrictions on
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Jr., Warren R.A.;
fimi: endoglucanase
                                                                                                                                                                                                                                                                                                                                                                             672
                                                                                                                                                                                                                                                     350 SLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDL 409
                                                                                                                                                                                                                                                                                                                                                 TKWDSGGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTS 469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28 SCSHVPDLAFNPWQEIALETDST-----FADIAFTEDP--NHGWLVGTKETIFETTD- 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -!- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  470 PVFTTGTSVDYAEL-----NPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -LHTTDGGQTWARIPLSEKLPGAPYSIIALGPQTAEMITDLGAIYKTTNGGKNW-KALVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WAASQG-----VPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GVMFHAVPGKEGDLWLAASSG-LYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WLLARGGOLOFSTDPDAEEWSDVIAPQ------DKG-----SWGLL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GGVTTGGTVAASADGSRFV-------WAPGDP------GQPVVYAVGFGNS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Micrococcineae; Cellulomonadaceae; Cellulomonas. NCBI_TaxID=1708;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
01-OCT-1996 (Rel. 34, Last annotation update)
Endoglucanase D precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
MEDLINE-9320993; PubMed-8458833;
Meinke A., Gilkes N.R., Kilburn D.G., Miller R.C. Jr., Warren R.
"Cellulose-binding polypeptides from Cellulomonas fimi: endogluc
D (CenD), a family A beta-1,4-glucanase.";
J. Bacteriol. 175:1910-1918(1993).
-:- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
                                                                                                                                                                                                             Indels 110;
                                                                                                                                                                Length 342;
EMBL; D90903; BAA17091.1; -.
InterPro; IPR002860; BNR.
Pfam; PF02012; BNR; 4.
Hypothetical protein; Complete proteome.
SEQUENCE 342 AA; 37291 MW; 9AB2A72B82DA046A CRC64;
                                                                                                                                                             3.6%; Score 145; DB 1; L 22.2%; Pred. No. 0.1; ative 48; Mismatches 130;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              747 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                             Query Match
Best Local Similarity
Matches 82; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  673 VVGTIGGVTG 682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LSPEKGFVLG 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Cellulomonas fimi.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HYDROLASES)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (Cellulase).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GUND_CELFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             315
                                                                                                                                                                                                                                                                                                                                              110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 233
                                                                                                                                                                                                                                                                                                                                                                                         78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             614
       DR DR DR SO
                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                         음
                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A PART OF THE PROPERTY OF THE
```

```
use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          97 GINVVRVPVSTQLLLEWKAGTFLKPNVNTYANPELEGKNSLQIFEYWLTLCQKYGIKVFL 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       405 ATNDLTKWDSGQIHIAPMVKGLEETAVNDLISP---PSGAPLIS----ALGDLGGFTH 456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  382 VLNP-NSGDTGGL--LLDDWKTWDEVKYSTMLEPTLWKHGGKYVGLDHQVPLGGVGSTTG 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  439 TSISQVGGGTPDT--TAPTAPTG------LRAGTPTASTVPLTWSASTDTGGS 483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                320 -----SWW------PDTIIFRSTD-GGATWTRIWDWISYPNRSLRYVL 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            356 DISAEPWLTFGVQPNPPVPSPKLG------WMDEAMAIDPFNSDRMLYGTGATLY 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            537 TARTAAGGGDVTAPSVPTGLTAGT------PTATSVPLTWTASTDTGGSGVTGY 584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        484 GVAGYEVYRGTTLVGTTTATSYTVTGLAADSAAYTFS-----VRAKDGAGNTSAASAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        328 DVWGPNWLFIHDEDIAPLLIGEWGGRLGQDPRQDKWM---AALRDLVAERRL---SQTFW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      157 DVHSAEADNSGHVYNMWWKGDITTEDVYEG-----WEWAAT--RWKDDDTIVGADIKNEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | | | : : :|| :: 210 HGTQGSTERAKWDGTTDKDNFKHFAETASKKILAINPNWLVFVEGVEIYPKPGVPWTSTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---GKNWFQGSEPGGVTTGG--TVAASADGSRFVWAPGDPGOPVVYAVGFGNSWAASQGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PA-NAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVAAGLP-----SSGAVGVMFH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          243 --HKGVFDPVNHVL---YIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              619 AVPGKEGDLWLAASSGLYHSTNGGSSWSAIT----GVSSAVNVGFGKSAP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      457 ADVTAV----PSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    252;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   179 DITGYQSDIQGVVWVAFDKSSSLGQASKTIFVGVADPNNPVFWSRDGGAT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     298 FGYSGL------TIDRQ-------HPNTIMVATQI------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Cellulose degradation; Hydrolase; Glycosidase; Repeat; Signal.
SIGNAL 1 39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ENDOGLOCANASE D.
FIBRONECTIN TYPE-III 1.
FIBRONECTIN TYPE-III 2.
PROTON DONOR (BY SIMILARITY).
NUCLEOPHILE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 747;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3.6%; Score 145; DB 1; Length 747
20.3%; Pred. No. 0.26;
.ive 69; Mismatches 234; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------WQA----VPGAPTGFIP--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   = ::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PS00659; GLYCOSYL_HYDROL_F5; 1.
                                                                                                                                                                                                                  InterPro; IPR001919; CBD_2.
InterPro; IPR003961; FN_III.
InterPro; IPR001362; FnIII_repeat.
InterPro; IPR001547; Glyco_hydro_F5.
Pfam; PF00159; CBD_2; I.
Pfam; PF00150; cellulase; I.
Pfam; PF00101; fn3; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRINTS; PRODO14; ENTYPEIII.
SMART; SMO0060; FN3; 2.
PROSITE; PSO0659; GLYCOSYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       78936
                                                                                                                                                         EMBL; L02544; AAA23089.1;
HSSP; P07986; 1EXG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity 20.3
41; Conservative
```

```
371
                                                              201
                                                                                                                                                                                                                                                                                                                                                                                517
                                                              g
                                                                                                                             ŏ
                                                                                                                                                                                     q
                                                                                                                                                                                                                                                   QΥ
                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FUNCTION: PLAYS A ROLE IN THE BIDIRECTIONAL TRANSPORT ACROSS THE NUCLEOPORIN COMPLEX (NPC). THE REPEAT DOMAIN HAS A DIRECT ROLE IN
   641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GFIPHKGVF--DPVNHVLYIATSNTGGPYDGSSGDVWKFS----VTSGTWTRISPVPSTD 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 293 TANDYFGYSGLTIDROHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLR 352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ---SSSSLGQ-----ASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPT 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        143 FGAP-SGKGLWRSTDSGATWSQMTNFPDVGTYI--ANPTDTTGYQSDIQGVVWVAFDK-- 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6 FGTPFGGTGGFGTT---STFGQNTGF---GTTSGGAFGTSAFGSSSNTGGLFGNSQTKPG 59
585 EV--YRGSTLVARPTGTSHTVTGLSAATAYTFTVRAVDAAGNVS-AASAPVGVTTAPDPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 3.6%; Score 143.5; DB 1; Length 937; Best Local Similarity 20.4%; Pred. No. 0.41; Matches 137; Conservative 78; Mismatches 222; Indels 233;
                                                                                                                                                                                                                                                                                                                                               01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
01-OCT-1996 (Rel. 34, Last annotation update)
Nuclear pore complex protein Nup98 (Nucleoporin Nup98) (98 kDa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Radu A., Moore M.S., Blobel G.;
"The peptide repeat domain of nucleoporin Nup98 functions as docking site in transport across the nuclear pore complex."; Cell 81:215-222(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       THE TRANSPORT.
SUBCELLULAR LOCATION: Nuclear pore complex.
DOMAIN: CONTAINS G-L-F-C REPEATS.
PTM: THE N-TERMINUS IS BLOCKED.
SIMILARITY: BELONGS TO THE GLFG FAMILY OF NUCLEOPORINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6B1F108C7390307A CRC64;
                                                           -GSSYPAVFVVGTIGGVTGAYRSDDCGTT----WVL 695
                                                                                                                         642 TGSCAVTYTANGWSGGFTAAVTLINTGTTALSGWTL 677
                                                                                                                                                                                                                                                                                       AA.
                                                                                                                                                                                                                                                                                       937
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLY-GLU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Interpro; IPR004325; Nucleoporin_FG.
Pfam; PF03093; Nucleoporin_FG; 26.
Nuclear protein; Transport; Repeat.
                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TISSUE-Liver;
MEDLINE-95254630; PubMed-7736573;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE 937 AA; 97809 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; L39991; AAC42054.1;
                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                nucleoporin)
                                                                                                                                                                                                                                                                             NU98_RAT
P49793;
                                                              665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             239
g
                                                                                                                                                                                                                                                                                       SO THE STATE OF TH
                                                              ò
                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
----LTFGVQPNPPVPSPKLGWMDEAMAIDPFN--SDRML 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---ELRLEDYQANRKGPQNQVGAGTTTGLFGSSPATSSATGLFSSSTTNSAFSYGQNKTA
                                                                                                                           397 YGTGATLYATNDLTKWDSGGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTH
                                                                                                                                                                                                                                                      457 ADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWF
                                                                                                                                                                                                                                                                                                                                                                                 QGSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                577 VNPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGL-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ------FGNKPTLTLGTNTNTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AYRSDDCGTTWVLI------NDDQHQYGNWGQAITGDHANLRRVYIGTNGRG
                                                                                                                                                                                                                                                                                                                  290 -QATTTPNTGFS----NTMGLFGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         463 AFGAPGFNTSTAILGFGAPQAPVALTDPNASAAQ----QAVLQQHLN-----S
                                                                                                                                                                                   258 FGTSTTGFGTN-----PGG-----LFGQQNQQTTSLFSKPFG------
353 YVLDISAEPW-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        729 IVYGDIGGAP 738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    507 LTYSPFGDSP
```

Search completed: July 2, 2002, 09:18:03 Job time: 371 sec

THIS PAGE BLANK (USPTO)

Title: Perfect score:

ü

Run

Seguence:

Scoring table:

Searched:

••

Database

```
093400 pseudomonas
0911pl bos taurus
0911bl bos taurus
0911bs caenorhabdi
093658 unidentifie
050379 mycobacteri
090710 acetivibrio
090711 lactococcus
090711 lactococcus
090731 mycobacteri
090731 mycobacteri
090311 meurospora
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN=A3(2);

MEDLINE=9700351;

Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,

Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,

Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,

Redenbach M., Rieser Gosmids and a detailed genetic and physical map for the B wb Streptomyces coelicolor A3(2) chromosome.";

Mol. Microbiol. 21:77-96(1996).

REDEL, ALOSTOSIS, CAA20642.1; -.

REDEL, PO7986, IEXG.

RESP, PO7986, IEXG.

RIEFPRO, IPR001319; CBD_2.

RIEFPRO, IPR001319; CBD_2.

Redem, PP002012; BNR, 10.

Redem, PP00203; CBD_2.

REDEL, AND PO2053; CBD_2.

REDEL, AND PO2053; CBD_2.
                              Q9p9a9 uncultured
Q98ln6 rhizobium l
Q9aaj5 caulobacter
                                                                               Q98mg7 rhizobium 1
O07219 mycobacteri
Q9n5k0 caenorhabdi
                                                                                                                                                                                                                                                                                                                                                                                              Q915n6 pseudomonas
Q99552 homo sapien
Q938a2 stigmatella
                                                                                                                                                                                                                                                                                                                                                                                                                                                 09a9p9 caulobacter
09hd43 homo sapien
018758 sus scrofa
                                                                                                                                                                                                                                                                                                                                                                               caenorhabdi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
NCBI_TaxID=1902;
                                                                                                                                                                                                                                                                                                                                                                               Q9nan8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN=A3(2);
Seeger K.J., Harris D.;
Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN=A3(2);
Parkhill J., Barrell B.G., Rajandream M.A.;
Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             890 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ALIGNMENTS
                                                                                                   007219
09N5K0
0934G0
093165
0093658
0003658
050379
                                                                                                                                                                                                                                                                      Q9CFL1
P91365
Q9F9U2
Q9F011
Q96U11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
                                                                                                                                                                                                                                                                                                                                                                                                Q915N6
Q99552
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9A9P9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9HD43
Q18758
                                                                                                                                                                                                                                                          09RM63
                                                                                                                                                                                                                                                                                                                                                                               Q9NAN8
                                                                                                                                  1023
                                                                                                                                                                                                                                                                        16
                                                                                                                                                                                                                                                                                                                           . 91
                                                                                                                                                                                                                                                                                                                                             3 .
                                                                                                                                                                                                                                                                                                                                                                                                16
                                                                                                                                                                                                                                                                                                                                                                                                                                                    16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-NOV-1998 (TrEMBLrel 08, 01-NOV-1998 (TrEMBLrel 08, 01-JUN-2001 (TrEMBLrel 17, PUTATIVE SECRETED CELLULASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                    631
1047
13055
1684
3716
1915
                                                                                                                                                                                                                                                                        1441
2232
695
2204
2117
3157
809
4180
3570
779
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Streptomyces coelicolor.
                                                                                 3145
699
1203
695
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                               167
166.5
166.5
166.5
164.5
164.5
163.5
163.5
163.5
163.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
156.5
15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             086727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
086727
Q92uu8 rhizobium m
P73139 synechocyst
P74440 synechocyst
Q9110 bos taurus
Q92411 bacillus sp
                                                                                                              2, 2002, 09:17:22 ; Search time 91.42 Seconds (without alignments) 1400.309 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Oghlq9 thermoplasm
Ogrk65 streptomyce
O38319 lactococcus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9wyel thermotoga
Q9i2m3 pseudomonas
Q91948 pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9aqhO caldiceliul
Q97kkO clostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9p4t8 agaricus bi
O74170 aspergillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  086727 streptomyce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                     1 ATTOPYTWSNVAİGGGGFVD......YIGTNGRGIVYGDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Description
                                Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                        Total number of hits satisfying chosen parameters:
            GenCore version 4.5
Copyright (c) 1993 - 2000 Comp
                                                                                                                                                                                                                                                                                                           562222 seqs, 172994929 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Listing first 45 summaries
                                                                                   OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q92UU8
P73139
P74440
Q9N1P0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q92411
Q9HLQ9
Q9RK65
Q38319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q9AQHO
Q97KKO
Q9P4T8
O74170
Q9WYE1
Q912M3
                                                                                                                                                                                                                                                      BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_organelle:*
sp_phage:*
sp_plant:*
sp_rodent:*
sp_virus:*
sp_virus:*
sp_virus:*
sp_virus:*
sp_virus:*
sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  086727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_human:*
sp_invertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Match 100%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sp_archea:*
sp_bacteria:*
sp_fungi:*
                                                                                                                                                                                                                                                                                                                                                                              Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                  Post-processing: Minimum Match 0%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16
16
                                                                                                                                                                             US-09-917-376-3
4036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SPTREMBL_19:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   856
707
2468
1800
2174
3972
4199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             998
848
1904
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Maximum
                                                                                                                   July
```

Score

Result . 일 2371 2062.5 2009 1753.5 1680 1160.5 238 202.5

189 186.5

185.5 185.5 184.5 182.5 182.5

extreme

D. P.

Williams

~

```
WGYNGVVSIAADPINTNKV-WAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGN 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTT 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYS 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4 QPYTWSNVAI-GGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNN 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              35 QPYVWKNVKIEGGGGFITGIVFNPKEKNLVYVRTDIGGAYRSTDGGNTWTQLMDWVSFDD 94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FCKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    302 GLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  362 WLTFG-VQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121 APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            481 AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQG-SEPGGVTTGGTVAASADGSRF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            540 VWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          F-QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., William Bergquist P.L.;
"Multidomain and multifunctional glycosyl hydrolases from the thermophile Caldicellulosiruptor isolate Tok7B.1.";
Curr. Microbiol. 40:333-340(2000).
                                                                                                                                                                                                                                                                                                                                                                                            Length 996;
                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                            51.1%; Score 2062.5; DB 2; 52.6%; Pred. No. 1.6e-107; ive 96; Mismatches 238;
                                                                                                                                                                                                                                                                                                                                    3C72B6ED22F3C614
                                                                                                                                                 InterPro; IPR002060; BNR.
InterPro; IPR001956; CBD_3.
InterPro; IPR001956; P_rich_extensn.
                MEDLINE=20171169; PubMed=10706665;
                                                                                                                                                                                                                                                                                                                                  108275 MW;
                                                                                                                                                                                                                                                    PRINTS; PRO1217; PRICHEXTENSN.
ProDom; PD001947; CBD_3; 1.
                                                                                                                                                                                                                                                                                                                                                                                            Query Match 51.1%
Best Local Similarity 52.6%
Matches 387; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RVYIGTNGRGIVYGDI 734
                                                                                                                                                                                                                                    Pfam; PF02012; BNR; 9.
                                                                                                                                                                                                                                                                                                                  966
                                                                                                                                                                                                                                                                                                                                      996 AA;
                                                                                                                                                                                                                                                                                             Hydrolase.
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                          PRINTS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        009
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      629
à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                            10;
                                                                                                                                                                                                                 185
                                                                                                                                                                                                                                                                                                                220
                                                                                                                                                                                                                                                                                                                                                       244
                                                                                                                                                                                                                                                                                                                                                                                            280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                688
                                                                                                                                       NGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGR 125
                                                                                                                                                                                                                                                                                                                                                                                                                                304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             457
                                                                              Gaps
                                                                                                                65
                                                                                                                                                                                                                                                                                               --EGAGYQYTTGFGTSWQASTGLPAGAIVESDRVNPAFFYGFKSGRFYVSTDGGATFTAS
                                                                                                                YTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY
                                                                                                                                                                                                                                                                        GMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQS
                                                                                                                                                                                                                                                                                                                                                       DIQGVVWVAFDKSS-SSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHK
                                                                                                                                                                                                                                                                                                                                                                                                                                GVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IDROHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKW-DSGGQIHIAPM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           484 NPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFVWAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AA-GLPSSGAVGVMFHAVPGKEGDLWLAASS-----GLYHSTNGGSSWSAITGVSSAVNV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GFGKAAPGASYQTLFTSAEIGGVRGIFRSTDAGATWTRVNDDAHQWGWTGAAITGDPRVY
                                                                            24;
                                  Length 890;
                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             09A0H0,
01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
GLYCOSYL HYDROLASE 5 (FRAGMENT).
Caldicellulosiruptor sp. Tok7B.1.
Bacteria; Firmicutes; Bacillus/Clostridium group;
                                                                            183;
                                  58.7%; Score 2371; DB 2; 57.7%; Pred. No. 7.8e-125; ive 109; Mismatches 183;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Thermoanaerobacter group; Caldicellulosiruptor.
NCBI_TaxID=80339;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                966
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RRVYIGTNGRGIVYGDI----GGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          749 GRVYVATNGRGVIYGDTSDTGGGTDPG 775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                         Conservative 109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A
                                                                         431;
                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09AQH0
                                                                       Matches
                                                                                                                                                                                                                                                                        126
                                                                                                                                                                                                                                                                                                                161
                                                                                                                                                                                                                                                                                                                                                  186
                                                                                                                                                                                                                                                                                                                                                                                         221
                                                                                                                                                                                                                                                                                                                                                                                                                                245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
0940H0
10 099
AC 099
AC 099
DT 071
DT 071
DT 072
DT 072
DT 073
DT 073
DT 073
DT 073
DT 074
DT
                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                        ò.
Ob
                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
```

420 450 480

ŝ

us-09-917-376-3.rspt

657

 $\sim$ 

a

```
GNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTD 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         bisporus.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60 WNNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLG 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Agaricus bisporus (Common mushroom).
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Agaricaceae; Agaricus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    628 FTQSSQTGLPTKGK-GI-FKTVIGHEGDIWIAGGKDGLWHSTDSGATFTKVSGVDASDTV
                                                                                                                                                                                                                                                                                                                       658 GFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 ATTQPYTWSNVAI-GGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVG
                                        449 SVKANGIEECAVNDVVVPTKGAQLLSAVGDDCGFYHDDITKVPSKMMTPNFSATTSIDY
                                                                                              481 AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQ-GSEPGGVTTGGTVAASADGSRF
                                                                                                                                                                       540 VWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVT
                                                                                                                                                                                                                                            600 F-QPVAAGLPSSGAVGVMFHAVPGKEGDLWLA-ASSGLYHSTNGGSSWSAITGVSSAVNV
                       APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            41;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 806;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Morales-Almora P., Thurston C.F.;
Morales-Almora P., Thurston C.F.;
Morales-Almora analysis of the cellulolytic genes in Agaricus I
Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AJ295929; CAC02964.1; -.
HSSP; P00725; 2CBH.
InterPro; IPR002860; BNR.
InterPro; IPR0020654; CBD_fungal.
PFam; PF02012; BNR; 8
PFmm; PF02012; BNR; 8
PFmm; PF02012; EMR; 8
PFMRT; SM00734; CBD_1; 1.
PROSITE; PS00562; CBD_FUNGAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          249; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                7C67E382E62FF341 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      / Match 1753.5; DB 3; Local Similarity 45.9%; Pred. No. 2.5e-90; les 341; Conservative 112; Mismatches 249;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               908
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CEL6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             23 PC
806 CE
84779 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                    746 GRVFVGTNGRGIVYGDIDGS 765
                                                                                                                                                                                                                                                                                                                                                                                              718 RRVYIGTNGRGIVYGDIGGA 737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-0CT-2000 (TrEMBLrel. 15, 01-0CT-2000 (TrEMBLrel. 15, 01-DEC-2001 (TrEMBLrel. 19, CEL6 PROTEIN PRECURSOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1
24 8
806 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=5341;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-D649;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Signal.
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               09P4T8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              09P4T8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09P4T8
                                                                                                                                .
qa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACCOCCOON DEATH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                       g
                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                              à
                                                                                                                                                                       ð
                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      241
                                                                                                                                                                                                                                                                                                                                                                     STRAIN-ATCC 824 / DSM 792 / VKM B-1787;
MEDLINE-21359325; PubMed-11466286;
Neolling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
Bennett G.N., Koonin E.V., Smith D.R.;
"Genome sequence and comparative analysis of the solvent-producing
J. Bacterium Clostridium acctobutylicum.";
J. Bacteriul. 183:4823-4838(2001).

EMBL. AROPGORS, AAK78895.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WLTFG-VQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNM 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTG 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYS 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3 TQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNN 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   183 YQSDIQGVVWVAFDKSSSSLGQASKTIFVGVAD-PNNPVFWSRDGGATWQAVPGAPTGFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEP
                                                                                                                                                                                                                                                Clostridium acetobutylicum.
Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
Clostridium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 839;
                                                                                                                                              01-OCT-2001 (TrEMBLrel. 18, Created)
01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
PROBABLY SECRETED SIALIDASE, SEVERAL ASP-BOXES AND DOCKERIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              90824 MW; 6C99A041CA9CF984 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
49.8%; Score 2009; DB 16;
Best Local Similarity 50.9%; Pred. No. 1.3e-104;
Matches 377; Conservative 107; Mismatches 240;
                                                                                                               AA.
                                                                                                               839
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF00404; Dockerin_1; 2.
Pfam; PF01344; Kelch; 1.
PROSITE; PS00018; EF_HAND; UNKNOWN_2.
                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002860; BNR.
InterPro; IPR002105; BNR.
InterPro; IPR002105; Bockerin_1.
InterPro; IPR002048; EF-hand.
InterPro; IPR001798; Kelch.
Pfam; PF02012; BNR; 9.
||::||||:|:|
| 745 RVFVATNGLGIKWGEI 760
                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Complete proteome.
SEQUENCE 839 AA;
                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=1488;
                                                                                                           Q97KK0
Q97KK0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     302
                                                                                              097KK0
```

윱 ò a ò g ò g ò a ò Q ŏ

10;

Gaps

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                             g
                                                                                                                                                                                             ò
                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              647
                                                                                                                                                                                                                                                                                                                                                                                                                                589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                655
GNMPGRGMGERLVVDPNLNSILYFGARSGNGLWKSTNSGRSWSKVTSFTDTGSFVPDPSD 201
                                                                                                                                                                                                                                                                                                                                                                                        434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Aspergillus aculeatus.
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
NCBI_TaxID=5053;
                                                                         GFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYF
                                                                                                                                                                                                                                             GYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDIS
                                                                                                                                                                                                                                                                                                                                        359 AEPWLTFGVQPNPPVPSP---KLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSG
                                                                                                                                                                                                                                                                                                                                                                 TSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ------VMVSRNQATFNIFYGASGKTFYVSTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              596 GGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              656 NVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYG-NWGQAITGDH
                                               TTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADP-NNPVFWSRDGCATWQAVPGAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Arai M., Takada G., Kawaguchi T., Sumitani J.;
Arai M., Takada G., Kawaguchi T., Sumitani J.;
Submitted G. Gun-1998) to the EMBL/GenBank/DDBJ databases.
EMBL; AB01551; BAA29031.1; -.
HSSP; P00725; 2CBH.
R InterPro; IPR002264; CBD_fungal.
R InterPro; IPR002244; CBD_fungal.
R Ffam; PF007134; CBD_1; 1.
R ProDom; PD001821; CBD_fungal; 1.
R PROSITE; PS00265; CBD_fungal; 1.
R PROSITE; PS00265; CBD_fungal; 1.
R PROSITE; PS00265; CBD_fungal; 1.
R SWART; SM00265; CBD_fungal; 1.
R SWART; SM00265; CBD_fungal; 1.
R PROSITE; PS00265; CBD_fungal; 1.
R PROSITE; PS00265; CBD_fungal; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    856 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RVYGRVYIGTNGRGIFYGDVAGS 730
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              715 ANLRRYYIGTNGRGIVYGDIGGA 737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                552 ADIILWRTNGNG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               08,
08,
17,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             074170 PRELIMINARY;
074170;
01-NOV-1998 (TEMBLEEL 0
01-NOV-1998 (TEMBLEEL 0
01-JUN-2001 (TEMBLEEL 1
AVICELASE III.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
  142
                                               180
                                                                                                                                             239
                                                                                                                                                                                                                                          299
                                                                                                                                                                                                                                                                                                                                                                                        379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           708
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
074170
                                                                                             g
                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AC
DTT DTT DTT DTT DTT DTT BD 
                                               ò
                                                                                                                                             ò
                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
```

```
17;
                                                                                                                                                                                                                                                                                                                                         SST--YTSDPVGIAWVIFDSTSGSGSATPRIFVGVADAGKSVFKSEDAGATWAWVSGE 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ---wnnwgyngvvsiaadpintnkvwaavgmytnswdpndgailrssdqgatwqitplpf 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                KLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIAN 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PT-GFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTAN 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DYFGYSGLTIDROHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVL 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    654 AVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQA--IT 711
                                                                                Gaps
                                                                                                                                                 1 ATTQPYTWSNVAI-GGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVG 59
                                                                                                                                                                                               21 AASQAYTWKNVVTGGGGGFTPGIVFNPSAKGVAYARTDIGGAYRLN-SDDTWTPLMDWVG 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |:|| ||::| | | |:|| | ||:|| | ||:|| ||:| ||:| ||: ||:| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                177 PIDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DISAEPWLTFGVQPNP-PVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    415 GGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     475 GTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              489 TNGIDYAGNKPSNIVRSGASD--DYPT---LALSSNFGSTWYADYAASTSTGTGAVALSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    595 DGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAI-TGVSS
                                                                                32;
Length 856;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Last sequence update)
Last annotation update)
                                                                            Mismatches 258;
   Score 1680; DB 3;
Pred. No. 3.4e-86;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ¥.
   41.6%; Score 1680;
46.2%; Pred. No. 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               712 GDHANLRRVYIGTNGRGIVYGDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    718 GDLQTYGRVFRGHERPGHLLRQSQREPAG 746
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Thermotoga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [1]
SEQUENCE FROM N.A.
STRAIN-MSB8 / DSM 3109;
MEDLINE-99287316; Pubmed=10360571;
                                                                            Conservative 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12,
12,
17,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bacteria; Thermotogales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1999 (TrEMBLrel.
01-NOV-1999 (TrEMBLrel.
01-JUN-2001 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ENDOGLUCANASE, PUTATIVE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Thermotoga maritima
                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=2336;
                                          Best Local Sim
Matches 346;
       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9WYE1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q9WYE1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                  09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         601
                                                                                                                                                                                                                                                                                                                                                                                                                                                117
```

S

```
26;
Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
Fyidence for lateral gene transfer between Archaea and Bacteria from
genome sequence of Thermotoga maritima.";
Nature 399:323-329(1999).
                                                                                                                                                                                                                                                                                                                                                                               MGVLSVALDPSDPKRIYAMTGKYTQDW-AGYGAILISEDYGETWTIVNLDKYGIKVGGNE 139
                                                                                                                                                                                                                                                                                                                                                                                                           182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              674
                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                               65
                                                                                                                                                                                                                                                                                               NGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLP - - - FKLGGNM
                                                                                                                                                                                                                                                                                                                                                                                                                                                               183 YQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             340 IK---DLNP------HWISD-VKIDPFDMNRAIFTTGYGVWVTYELKKSFEG----MGK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQ-GSEPGGVTTGG--TVAAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TFYRSTDGGVTFQPVAAGLPSSGAVGVMFHA----VPGKEGDLWLAAS-SGLYHSTNGGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSOMTNFPDVGTYIANPTDTTG
                                                                                                                                                                                                                                                                                                                                                                                                                         HKGVFDPVNHVLY1ATSNTGGPYDGSSGDVWKFSVTSGTWTR1SPVPSTDTANDYFGYSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MVK-----GLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPA---NAQIRSDRVNPKTFYAL--SNG
                                                                                                                                                                                                                                       Query Match 28.8%; Score 1160.5; DB 16; Length 707; Best Local Similarity 35.2%; Pred. No. 3.2e-57; Matches 264; Conservative 123; Mismatches 275; Indels 89;
                                                                                                                                                                                                  8CD8743CBDA6A99F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          79496 MW;
                                                                                                                                                      InterPro; IPR002860; BNR. Pfam; PF02012; BNR; 9.
                                                                                                                                                                                               707 AA;
                                                                                                                                                                                   Complete proteome.
                                                                                                                                             TM0305;
                                                                                                                                                                                                SEQUENCE
                                                                                                                                            TIGE;
                                                                                                                                                                                                                                                                                                                                                   99
                                                                                                                                                                                                                                                                                                                                                                               81
                                                                                                                                                                                                                                                                                                                                                                                                         123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        704
                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
    RA
RA
RA
RA
RA
DR
CR
CW
KW
KW
KW
KW
                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ô
```

```
1274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WT-----SYPNRSLRYVLDISAEPWL--TFGVQPNPPVPSPKLGWMDEAMAIDPFNS 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           393 DRMLYGT----GATLYATNDLTKWDSGGQIHIAPMVKGLEETAVNDLISP----PSGAPLI 445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                SEQUENCE FROM N.A.

SEQUENCE FROM N.A.

MEDLINE-20437377; PubMed-10984043;
Stover C.K., Pham X.-Our., Erwin A.L., Mizoguchi S.D., Warrener P., Stover C.K., Darinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou Mickey M.J., Brinkman F.S.L., Hufnagle W.O., Rowalik D.J., Lagrou M. Carber K.L., Goltury L., Tolentino E., Westbrock Wandnan S., Yuan Y., Bridy L.L., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Complete genome sequence of Pseudomonas aeruginosa PAOI, an opportunistic pathogen.";

Nature 406:959-964(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1188 -DAGGNPIGOVTAD-----GTPLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1215 NGTVIVATATDPTGRTGPQAATTVDAVAPPAPVIDPSNGTTISGTAEAGAKVILTDGNGN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HVLYIATSNTGGPY - - - DGSSGDVWKFSVT - SGTW - - TRISPVPS - - - - TDTANDYFGYS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SLPATTTVDSSLPSIPQVDPSNGSVISGTADAGNTIII --TDGNGNPIGQVTADGSGNWS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1084 VNLSNGSSLSG----TAEPGSTVILTD-----GNGNPIAEVTADGSGNWTYTPSTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --NG-VVSIAADPINTNKVWAAVGMYTNS------WDPNDGAILR-SSDQGATWQITPLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             116 FKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSOMTNFPDVGTYIA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N------PTDTTGYQS----DIQGVVWVAFDKSSSSL--GQASKTIFVGVADPN-N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PV-FWSRDGGATWQAVPGAP------TG-----FIPHKGVFDPVN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1275 PIGETTADGSGNWSFTPGTPLANGTVVNAVAQDPAGNTGPQGSTTVDAVAPNTPVVNPSN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GNLLNGTAEPGSTVTLTDGNGNPIGQTTADGSGNWSFTPGSQLPNGTVVNVTASDAAGNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GL----TIDRQHPNTIMV----ATQISWWPD---TIIFRSTDGG-----ATWTRIWD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11 VAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels 292;
                                                                                                                                             Proteobacteria; gamma subdivision; Pseudomonadaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 2468;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 protein; Complete proteome.
2468 AA; 238414 MW; 13596AFAB2C4B899 CRC64;
                                                     Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 238; DB 16;
Pred. No. 6.3e-05;
88; Mismatches 314;
                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; AE004613; AAG05263.1; -.
InterPro; IPR001343; Hemlysn.Ca_bind.
InterPro; IPR0013480; Phosphopant_attach.
PRINTS; PR00313; CABDNAGRPT.
PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FTPGIPLPDGTVVNVVARSPSNVDSAPAVITVDGVAPAAPV
 AA
                                    Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5.9%;
                                (TrEMBLrel. 16, TrEMBLrel. 16, TrEMBLrel. 18,
                                                                                       HYPOTHETICAL PROTEIN PA1874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 22.8
Matches 205; Conservative
                                                                                                                              Pseudomonas aeruginosa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                 NCBI_TaxID=287;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hypothetical
                                                     01-MAR-2001
01-OCT-2001
                                                                                                                                                   Bacteria;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Q912M3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       302
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

50;

```
72 AADPINTNKVWAAVGMYTN-SWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGER 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              :||:
1583 FGDL 1586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                731 YGDI 734
                           908 KAD-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q920U8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
Q92UU8
                          a
                                                                                                                            g
                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                    ö
                                                                                                    Qγ
                                                                                                                                                      δ
                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                        δy
                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      41;
1605 GQ-----VTADGSGNWSFTPGTPLANGSVINALAQDAAGNNSSPTSATVDSLAPAAPV 1657
                                                                                                                                                                                                                                                                 1760 LTDGNGNPIGOTLADGSGNWSFTPGTPLANGTVVNAVAQDPAGNTSGPASTTVDTVAPAT 1819
                                                              1549 AVAQDPAGNTSGPASVTVDAIAPPAPVINPSNGVVISGT----AEAGATVILTDGNGNPI 1604
                                                                                                                                                                                                                 ----LSNGTVVNAVAQDAAGNTSGPVSTTVDAVAPATPVIDPSNGVELSGTAEPGVRVI 1759
                                                                                                                                                                                        KTFYALSNGTFYR-----STDGGV--TFQPVAAGLP---SSGAVGVMFHAVPGKEGD 626
                                                                                                                                                                                                                                           681
                                     SALGDLGGFTH-----ADVTAVPSTIF--TSPVFTTGTSVDYAELNPSIIVRAGSFDPS 497
                                                                                                                                                                                                                                                                                                                   ------PVINPSN-----GSVITGT-AEVGAKVILTDGNGNPIGETTADGSG 1859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            12 AIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSI 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
MEDLINE-20225830; Pubmed-10762233;
Septionsa-Urgel M., Salido A., Ramos J.L.;
"Genetic analysis of functions involved in adhesion of pseudomonas putida to seeds.";
                                                                                                                                        ----TGGTVAASADGSRFVWAPGDPGQPV--VYAVGFGNSWAASQGVPANAQIRSDRVNP
                                                                                      ------QGSEPGGVT------
                                                                                                                                                       LWLAASSGLYHS-TNGGSSWSAITGV----SSAVNVGFGKSAPGSSYPAVFVVGTIGGVT
                                                                                                                                                                                                                                                                                            GAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDIGGAPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 5.3%; Score 214; DB 2; Length 1800; Best Local Similarity 22.0%; Pred. No. 0.00093; Matches 186; Conservative 107; Mismatches 319; Indels 232;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       185196 MW; 409BFC017552F96E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                       01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
CALCIUM-BINDING OUTER MEMBRANE-LIKE PROTEIN (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ς.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROSITE; PS00330; HEMOLYSIN_CALCIUM; PROSITE; PS50234; VWFA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                           Created)
                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bacteriol. 182:2363-2369(2000).
                                                                                      498 SQPNDRHVAFSTDGGKNWF-----
                                                                                                                                                                                                                                                                                                                                                                                                        01-OCT-2000 (TrEMBLrel. 15, 01-OCT-2000 (TrEMBLrel. 15,
                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1800 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pseudomonas putida
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=303;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                               Q9L948
Q9L948;
                                     446
                                                                                                                                        526
                                                                                                                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                                                                                                                                                                                                                                      Q9L948
             g
                                                            QQ
                                                                                                            g
                                                                                                                                                          g
                                                                                                                                                                                       ö
                                                                                                                                                                                                                염
                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                 g
                                     ŏ
                                                                                     ò
                                                                                                                                        ò
                                                                                                                                                                                                                                          å
                                                                                                                                                                                                                                                                                            à
```

```
| : :| | : | | : | | : | | : | | : | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | : | | : | | : | | : | | : | | : | | : | | : | : | | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1090 AVTDLSSSNVKMYQNVTEMANAGLGVSDLHTVNGQ------SYYDGYKLNEGPEG 1138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1194 LSSLTLTPPAYYKGSFDITVTSTATESLGGSAITTGNIPVTVYGATYKASVGTSGNDTLT 1253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1254 GSEGNDIFVAD-VSGLNVVQGKNY----NIAFMVDSSGSMSDKSIADAKTQLASVFNTLK 1308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1421 DVVNVNNYKMGDTFSAWADATH------KVEISSSGVVKVLTY-TENRRGELV 1466
                                                                                                                                                                                                                                                                                                                                                              --IFVGVADPNNPVFWSRD 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   382 DEAMAIDP-----NDLTK 411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       412 WDSGGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPV 471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        575 DRVNPK-----TFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLW 628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      686 SDDCGTTWVLINDDQHQYGNWGQAITG-DHANLRRVYIG-------TNGRGIV 730
                                                                                                                                                                                                    131 LAVDPNNDN-ILYFGAP--SGKGL----WRSTDSGATWSQMTNFPDVGTYIAN-----P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            226 GG----ATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      281 TWTRISP-----VPSTDTANDYFGYSGL-TIDRQHPNTIMVATQISWWPDTIIFRSTDG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                334 G-----ATWIRINDWISYPNRSLR-----YVLDISAEPWLIFGVQPNPPVPSPKLGWM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    520 -EPGGVTT----GGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1365 YEDAFKTTSNFFNSTMATSNKGAENLTYFITDGKPTYYO----SNESTNPSLWKNGKSLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      629 LAASS--GLYHSTNGGS-SWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 472 FTTGTS------VDY-AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGS--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                plasmid bymb (megaplasmid 2).

Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
Rhizobiaceae; Sinorhizobium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
HYPOTHETICAL GLYCINE-RICH PROTEIN SMB1548.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2174 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Rhizobium meliloti (Sinorhizobium meliloti)
----YAQFKFKPNDGTNL-
                                                                                                                                                                                                                                                                                                                                                        178 TDTTGYQSDIQGVVWVAFDKSSSSLGQASKT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           [1]
SEQUENCE FROM N.A.
```

```
Bacteria; Cyanoba
NCBI_TaxID=1148;
                                                                                                                        SLR1028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    94
     ACCOOR NAME OF THE PROPERTY OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                              36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1544 KSGNVTVSLDAGRTIQASGKDSIGIFAQSDAGTDNNCTIDVTVNGTVTGGSGDNGAGIWV 1603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          976 SIGGGG-------GNGCVGSVNSKEIGSGFNLTA-------NVGVGGSGS 1011
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1012 GGASGGN----AIVGL-----DSGTHLQTSGSGARGVIVQSIGGGGTSQGASVGLS 1059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            71 IAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGA------TWQITPLPFK 117
                                                                                                                The complete sequence of the 1,683-kb psymB megaplasmid from the N2-fixing endosymbiont Sinorhizobium meliloti.";
Proc. Nall. Acad. Sci. U. S.A. 98:9889-9894(2001).
EMBL; AL603645; CAC49389.1; -
Plasmid; Hypothetical protein; Complete proteome.
SEQUENCE 2174 AA; 203314 MW; 008EB68297B44182 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                             70
                  MEDLINE-21396508; PubMed-11481431;
Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
Vorhoelter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
Golding B., PubHier A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1171 CFGVSVGAT---IDDGGTGGTAANGNAVTLTHAGHIATAGDWADGIVAQSIGG--GGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | : | | : | | : | | : | | CTSTASGSQATANITVGVGGSGGAGGAVGITFDDNHGNSISTAGYSAY---GVLLQS
                                                                                                                                                                                                                                                                                                                                                                                                                           AIGGGGFYDGIVFNEGAPGILYVRT-DIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             118 LGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----GATWQAVPGAPTGFIPHKGVFDPVNHVLYIATS------NTGGPYDGSSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTIIFRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GAVIVDLISQSSI----TTQGIAAHGLVAQSIGGGGGVGGAASGAPLSFTGNSPGSYGDGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ASASLPGGG-EEAAEAEESEEGSGAFSASVGVSVGRTGGSGGS-SGTVNVTTAGT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---ISAEPWLTFGVQPNPPVPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PKLGWMDEAMAI - - DPFNSDRMLYGTGATLYA - - - - - TNDLTKWDSGGQIHIAPMVKGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EETAVNDLISPPSGAPLISAL - - - - - GDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  586 SNGTFYRSTDGGVTFQPVA----LWL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1283 IGGGGG----QGGDGSDEAAGRITVGGGFGGSGGAGGSGGMVTAKGWINLSTSG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           481 AELNPSIIVRAGSFDPSSQPNDRH--VAFSTDGGKN------WFQGSEPGGVTTGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TVAASADGSRFV - - - WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TDTTGYQSDIQGVVWVAFDKS---SSSLGQASKTIFVGVADPNNPVFWSRDG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 2174;
                                                                                                                                                                                                                                                                                                                                                                              197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTI-GGVTGA
                                                                                                                                                                                                                                                                                                                       5.0%; Score 202.5; DB 16; Length llarity 22.7%; Pred. No. 0.0052; Conservative 90; Mismatches 312; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TDGGATWTRIWDWTSYPNRSLRYVLD------
                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
Matches 176; Conserv
                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 10
P73139
ID P73139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1437
                                                                                                                                                                                                                                                                                                                                                                                                                             12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 227
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

3972 AA.

PRT;

PRELIMINARY;

P73139

```
42;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2458 VNGDGFADVISGGSPAGGVLIFGNSTKDLL--DAALGTDDLIISVENAQVKEFVALGDFD 2515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2859 SATVVPGLDGDTLAVA-----FRSNNSPATWVGLLNSSDVTNWQGSAELTQVDA 2907
                                                                                                                                                                                                                                                                                                                                                                                                                   ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----NDNIL/YFGAPSGKGLWRST------DSG-ATWSQMTNFPDVGTYIANP 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            "Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions."; DNA Res. 3:109-136(1996).
EMBL; D90904; BAA17165.1; -.
HSSP; P30617; LJPC.
                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
MEDILTRE=97061201; PubMed=8905231;
Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
Kineko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura T.,
Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
Tabata S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2634 PVPNIGGFGGGLQYFTYEAGNAVLQPTVNPPNASVTEASGLSSWGQI-SFPNQYAQAGVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2693 SFATLDGWLYQAFYGINERISTKDSYIYIQRSRDGVSWENLTQVVPLDSNGTPIDLKNLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              245 GVF----DPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRI------WD------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----RWIPLLDW-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---VGWNNWGYNG---VVSIAADPINTNKVWAAVGMYTNSWD--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2516 GDGLADFGVIDDQGNFFLVLGSPELGSQGSLVIDSTLPNLSNFNQAWGVGDFNGNGYDDF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---GRGMGERLAVDPN--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               178 TDTT -- GY -----VAFDKSSS -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       201 -SLGQASKTIFVGVADPNNPVFWSRDG------GATWQAVP-----GAPT--GFIPHK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4.7%; Score 189; DB 16; Length 3972;
19.3%; Pred. No. 0.063;
Live 93; Mismatches 276; Indels 326;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SMART; SM00108; B_lectin; 1.
SMART; SM00237; Calx_beta; 1.
SMART; SM00191; Int_alpha; 9.
PROSITE: PS00284; SERPIN; UNKNOWN_1.
Hypothetical protein; Complete proteome.
SEQUENCE 3972 AA; 418281 MW; E9FE49F274FBEABD CRC64;
                                                                                                                                                                           Cyanobacteria; Chroococcales; Synechocystis
01 FEB-1997 (TrEMBLrel. 02, Last sequence update) 01-JUN-2001 (TrEMBLrel. 17, Last annotation update) HYPOTHETICAL 418.3 KDA PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13 IGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ----PNDGAILRSSDQGATWQITPLPFKLGGNMP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            93; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR001480; B_lectin.
InterPro; IPR003644; Calx_beta.
InterPro; IPR000413; Integrin_alpha.
InterPro; IPR000215; Serpin.
Pfam; PF01453; Agglutinin; 1.
Pfam; PF01839; FG-GAP; 8.
                                                                                                                                       Synechocystis sp. (strain PCC 6803)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRINTS; PR01185; INTEGRINA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 4.73
Best Local Similarity 19.33
Matches 166; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
```

```
3195 GVWTQFHDDG 3204
  2261
                                                 δ
                                                                                             8
                                                                                                                                                    δ
                                                                                                                                                                                        염
                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                               οy
                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                 δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       54;
                                                                                                                                         GGVASILFNQSFILSLNGSNNESLLFAFSNSLFEPNQASRW--GEQVRDIGDFNGDGIAD 3014
                                                                                                                                                                                                                                                                                                                                                                                                                                  3188 APGNFDNLSYVLFGSDFTNQVNQLGTIGDDVMLGSPTGEIFVAGQGDDQIYTNGGVDTVY 3247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŝ
  -WISYPNRSLRYVL-----DISAEPWLTFGVQPNPPVPSPKL 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y., Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T., Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S. Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M., Tabata S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions.";
                                                                                                                                                                                                                                                                                                                               ---IAPMVKGLEETAVNDLISPPSGAPLISALGDL----GGFTHADVTAVPSTIFTSPVF
                                                                                                                                                                                                                                                 3015 LAVLAPGYRNLLQFPILDY-----PAINNLGGVFIYYGEESGISVNDPPDVVLAAPDL
                                                                                                                                                                                                                                                                                                   TTGTSVDYAELNPSIIVRAGSFDP--SSQP------NDRHVAFSTDGGKNW--
                                                                                                                                                                                                                                                                                                                                                                                                -FQGSEP-----GGVTTG-----GTVAASADG-SRF-VW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   542 APGD----VPANAQIRSDR
                                                                                                GWMDE-----AMAIDPFNSDRMLYGTGATLYATNDLTKWDSGQIH------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    389;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 4199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hypothetical protein; Čomplete proteome.
SEQUENCE 4199 AA; 442441 MW; 1DB4BC556D2C1B5C CRC64;
                                              2908 NSQVSLTVVDGTYYLFFTSSTEASASYATSTDGLNWGDITLIPW----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synechocystis sp. (strain PCC 6803).
Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
NCBL_TaxID=1148;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              P74440;
01-FBB-1997 (TrEMBLrel. 02, Created)
01-FBB-1997 (TrEMBLrel. 02, Last sequence update)
01-FDN-2001 (TrEMBLrel. 17, Last annotation updat
HYPOTHETICAL 442.4 KDA PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4.6%; Score 186.5; DB 16; larity 19.2%; Pred. No. 0.093; Conservative 124; Mismatches 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DNA Res. 3:109-136(1996).

EMBL; D90915; BAA18541.1; -.

InterPro; IPR002860; BNR.

InterPro; IPR003644; Calx_beta.

InterPro; IPR000413; Integrin_alpha.

InterPro; IPR002989; Mycobact_pentapep.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF02012; BNR; 3.
Pfam; PF01839; FG-GAP; 5.
SMART; SM00237; Calx_beta; 1.
SWART; SM00191; Int_alpha; 8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A. MEDLINE-97061201; PubMed-8905231;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3248 AGPGNDFVTVTDTNFRRLDGG 3268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                577 VNPKTFYALSNGTFYRSTDGG 597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Simi
Matches 198;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SLR0408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            P74440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           11
343
                                                                                                379
                                                                                                                                                 2957
                                                                                                                                                                                                420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT
P74440
ò
                                              엄
                                                                                                ò
                                                                                                                                            g
                                                                                                                                                                                           δ
                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACCOORDING TO THE PROPERTY OF ```

```
ADVASDNGFVIDGNLIGNPPTTFETTSQYIDTTPAILINGSNLYLAYKGFGGNNQ---IX 2317
                                                                                                                                                                                                                                       375 TWN----APLALGGISSTPPTLFVYQGTLSLFAANNSTSTVLQFYLNSSNEWIYANEIGS 2431
                                                                                                                                                                                                                                                                                                                                                                2432 NQTAISAISATVLGDTLYLVYKGGTRNTPSTLDYITSTTN----ADLSANDWSSIPIPG 2486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3100 NGSTWV--NNCPYQ----GDGWRSAITQMA----VQWGED--GS--PSQIVVGLADGAVI 3145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2929 LRANQSQGSNPRGFAIDGSPNSQAGIALNGGGDINGDGFADFIIGAPGENNLQYNQQIVF 2988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1989 IENGELSDDDKYSYILYLDGNQTIQMGGGDWQANQ-VWTN-QVATNWNNSSRPPEAVIGQ 3046
                                                                                                                                     FTVSTNNGQSWNSEVQL---PQSAQTIFPPAIAFFNNVLYLAYVDGNNGLNIITSQDQGQ 2374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TWQITPLPFKLGGNMPG-----RG-MGERLAVDPNNDNILYFGAPSGKGLWRSTDSGA 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPST------DTAN 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   683 AYRSDDCGTTWVLINDDQHQYGNWGQAITG--------DHANLRRVYIGTN 725
                                                                                DWVGWNN-WGYNGVVSIAADPINTNKVW-AAVGMYTNSW-----DPNDGA-ILRSSDQGA 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                            160 TWSQMT------NFPDVGTYIANPTDTTGYQSDIQGVVWVAF---
                                                                                                                                                                                                                                                                                                                                                                                                                         --DKSSSSLGQASKTIFVGVADPNNPV-FWSRDGGATWQAVPGAPTGFI---PHKGVFDP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DYFG--YSGLTIDRQHPNTIM-----VATQISWWPDTIIFRSTDGGATWTRIWDWTSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DGFADVFSGGT----NAGAIIFGNSTKDLLTTASGSEDLVI-----SV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PNRSLRYVL---DISAEPWLTFGV------QPNPPVPSPKL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----SGGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2749 SDGIPEIAIGSDERKIAGQISTSGSFSLLPTPTTSSVINTLAAANQLENIGDFNGDGIAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VTAVPSTIFTS-----PSIIVRAGSFDPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2869 GQISTYQLSRIAQAGDVNGDGFDDLLISSPYTVDAENNQGGVFVVFGGDDWNNQPFDLGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----DPGQPVVYAVG-----FGNSWAASQGVPANAQIRSDRVN----PKTFYAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SNGTFYRSTDGGVTFQ------PVAAGLPSSGAVGVMFHAVPGK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----FQGSEPGGVTTGGTVAASA------DGSRF----VWAPG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3047 SNGDIWYYPGGNQNWQSWGKLPAEINELAVNWNTSGNPQIIAGLGGKG--GIEYY----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  624 EGDLWLAASSGLYHSTNGGSSW-SAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----PNDRHVAFSTDGGKNW----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GWMDEAMAIDPFNSDR----MLYGT-----GATLYATNDLTKWD-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                726 GRGIVYGDIG 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SQ----
                                                                                                                                     2318
                                                                                                                                                                                                                                                                                                                                                                                                                      196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        498
                                                                                                                                                                                             108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             414
                                                                                26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      516
```

RESULT

SNVAIGGGGFVDGIV------FNEGAPGILYVRTDIGGMYRWDAANGRWIPLL

σ

ò

```
SPGVTRITALRGSETRVPSTGVSGLPGSTQGG---SAATGGSGA---GSGPTAPVSGETR 680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  120 GNMPGRGMGERLAVDPNNDNILYFGA-----PSGKGLWRSTDSGATWSQMTNFPDVGTY 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     228 ATWQAVPGA----PT-----GFIPHKGVFDPVNHVL--YIATSNTG---G 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      264 PYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWP 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       324 DIIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE-PWLIFGVQPNPPVPSPKLGWMD 382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               441 GATGTDVLRS-----GTSLPV-----SGGAVTPAPSPGGSSATA-----GPGVGS 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  481 ATTVQASGATGADVLRSGTSLPVSGVAVSPGSSPG-------RSGATAVSS-- 524
                                                                                                                                                                                                                                                                                                                                                    over 50
                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
STRAIN=BREED ANGUS;
MEDLINE-20232353; PubMed-10759843;
MEDLINE-20232353; PubMed-10759843;
Jiang W., Gupta D., Gallagher D., Davis S., Bhavanandan V.P.;
"The central domain of bovine submaxillary mucin consists of over 5 trandem repeats of 329 amino acids: chromosomal localization of the BSMI gene and relations to ovine and porcine counterparts.";
EMBL, AE178428; AAF67279.1; -.
                                                                                                                     Bos taurus (Bovine).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
Bovidae; Bovinae; Bos.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   569 TPGVDSQQTASLPAAA--RPTALGPGTSAPSGETSESRSSVPGGSETTQQPGAGSESPTL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AVGV-MFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           442 APLISALGDLGGFTHADVTAVP-STIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              558 NSWAASQ---GVPANAQIRSDRVNPKTFYALSNGTFYRST-DGG--VTFQPVAAGLPSSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              226 -ATPGSTTGRAA---GAGTPAVDSQQTASLPAAARPTALGPGTSAPSGETSESRSSVPGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             282 SETTQQPGAGSESPTLSPGVTRTTALRGSETRVPSTGVSGLPGSTQGGSAATGGSGAGSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              385 GSETASPLSGAAGTSATGSGTSIPPSG----APVTPEPPLISTGASAGPPASSESTVTLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           501 NDRHVAFSTDGGKNWFQGSEPGGVTTGGT-VAASADGSRFVWA--PGDPGQPVVYAVGFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     174 IANPIDITGYQSDIQGVVWVAFD-KSSSSLGQASK--TIFVGVADPNNPVFWSRD---GG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGQIHIAPMVKGLEETAVNDLISPPSG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels 151;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              818
73014 MW; B5BB44F84F66F86B CRC64;
                                     01-OCT-2000 (TrEMBLrel. 15, Created)
1-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
SUBMAXILLARY MUCIN (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 4.6%; Score 185.5; DB 6; Best Local Similarity 25.9%; Pred. No. 0.014; Matches 159; Conservative 55; Mismatches 249;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      671 VFVV~GTIGGVTGA 683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  818 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             \begin{array}{c} 1 \\ 818 \end{array}
                                                                                                                                                                                                                                                                                                                                                                                                                                                          NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
MEDLINE-21129642; PubMed-11234960;
Pastor F.I.J., Dylol X., Blanco A., Vidal T., Torres A.L., Diaz P.;
PMolecular cloning and characterization of a multidomain endoglucanase
from Paenibacilius sp BP-23: evaluation of its performance in pulp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; : | | | : | | 355 WGSLRYAANAAFISFVYSDWVKDPKKSRY--ONFATSQINYILGDNPRQSSYVVGYGQN 412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                79 NKVWAAVGMYTNS---WDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDP 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              147 NELWGQVGAGNTDHAWWGPAE------VMQMNRPSFKIDASCPGSDLAAETAAAL 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        196 AASSIVF------ADSDPVYSAKLLQHAKELYNFADTYRGKYTDCITDAAAFYN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           283 TRISPVPSTDTANDYF - - GYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----KYTPGGLAWLDQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                136 NNDNILYFGAPSGKGLWRSTDSGATWS-----QMTNFPDV--GTYIANPTDT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 --TGYQSDIQ-GVVWV------AFDKSSSSLGQASKTIFVGVADPNNPVFW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       23 SRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         293 TQG----WDS-----KHYG----AQILLARITSNLNMP-----EATKF----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               341 W------DWTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLG----
                                                                                                                                                                                                                                                                                                                                                                                                         refining.";
Appl. Microbiol. Biotechnol. 55:61-68(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          997 AA; 106927 MW; F20CB119D6410629 CRC64;
                                                                                                                             01-MAY-1999 (TrEMBLrel. 10, Created)
01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
CELLULASE PRECURSOR (EC 3.2.1.4).
                                                                                                                                                                                                                   Bacillus sp. BP-23.
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              248;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4.6%; Score 185.5; DB
21.1%; Pred. No. 0.018;
tive 85; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PROSITE; PS00592; GLYCOSYL, HYDROL, F9_1; PROSITE; PS00698; GLYCOSYL, HYDROL, F9_2; PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001956; CBD_3.
InterPro; IPR001956; FN_III.
InterPro; IPR001701; Glyco_hydro_9.
InterPro; IPR000566; Lipocln_cytFABP.
Pfam; PF000942; CBD_3; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF00759; Glyco_hydro_9; 1.
ProDom; PD001947; CBD_3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                             Appl. Microbiol. Biotechnol. 55
-!- SMILARUTX: CONTAINS 2 FIBE
--- EMBL; AJ133614; CAB38941.1; --
HSSP; P26221; 11F4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Glycosidase; Hydrolase; Signal SIGNAL 1 34 Po
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 21.19
Matches 165; Conservative
                                                                                                  PRELIMINARY;
SMART; SM00060; FN3;
                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                     NCBI_TaxID=89769;
                                                                                                                                                                                                                                                                                                                           STRAIN-BP-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                  Q924I1
                                                                RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                      염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

43;

```
686 SDDCGTTWVLINDDQHQYGN 705
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-MAY 2000 (TrEMBLEEL 13, 01-MAY 2000 (TrEMBLEEL 13, 01-JUN-2001 (TrEMBLEEL 17, PUTATIVE SECRETED PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            607 LP----SSG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SCF11.25.
Streptomyces coelicolor
                                                                                                                                                                                                                                                                                                                                              554 NIT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09RK65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 15
                              290
                              qq
                                                                                                                        δ
                                                                                                                                                       a
                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                  δλ
                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                      οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                           Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                47;
                        413 SPQHPHHRTAHSSWMNNEDI--PANHRHILYGAMVGGPNASDQYTDDIGDYVSNEVATDY 470
                                                             441
                                                                                         530
                                                                                                                          492
                                                                                                                                                                                                          | : | | | : | | 387 ITANFSNTKIYPGGEGNYRKEVQFRITGPQGAWNPANDPSYQNLTTGNPVKSNYIPVYDA 646
                                                                                                                                                                                                                                                  580
                                                                                                                                                                                                                                                                                                                                                                           632 SSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFV-----VGTIGGVTG 682
 ---TKW 412
                                                                                                                                                                                                                                                                                GVKVSGQEPGVTPVAVPAAPAG--VQAVA-GNSQVALNWSASAGAVSYTVKRAEVSGGPY 703
                                                                                                                                                                                                                                                                                                               TFYALS-NGTFYRST--DGGVTFQPVAAGL-----PSSGAVGVMFHAVPGKEGDLWLAA 631
                                                                                                                                                                                                                                                                                                                                                                                                          817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        54 -LLDWVGWNNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQIT 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     188 ANALHGNGKLVPGVFYYDIGPTFNISFPFSLKLYLNSSLVGGNDAIYFNYSIENGKITKS 247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-DSM 1728;
MEDLINE-20479972; PubMed=11029001;
Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
"The genome sequence of the thermoacidophilic scavenger Thermoplasma acidophilum.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9 SNVAIGGGGFVDGIVFNEGAPGI-----LYVRTD-IGG----MYRWDAANGRWIP- 53
                                                            DSG------ETAVNDLISPPSG
                                                                                         471 NAGFTGALAKMNLLYGONHOPLANFPAPEVKGDEYFVEAAVRSSGSNYTEIRALLNNRSG
                                                                                                                          442 APLISALGDLGGFTH----ADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAG----
                                                                                                                                                    531 WP--ARMGDQLSFKYFLDLSEVYAAGRIV--SDVQVTVSSSEGATVSQPVVVDAAKRIYA
                                                                                                                                                                                                                                                                                                                                 -----SFDPSSQPNDR-HVAFSTDGGKN-WFQGSEPG--GVTTGG------
                                                                                                                                                                                                                                                  --TVAASADGSRFVWAPGDPGQPVVYAVGFGNS-----WAASQGVPANAQIRSD-RVNPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              289;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Thermoplasma acidophilum.
Archaea; Euryarchaeota; Thermoplasmales; Thermoplasmaceae;
Thermoplasma.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
4.6%; Score 184.5; DB 17; Length
Best Local Similarity 20.1%; Pred. No. 0.02;
Matches 173; Conservative 113; Mismatches 285; Indels
----WMDEAMAIDPFNSDRMLYGT----GATLYATNDL--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Complete proteome. 
SEQUENCE 998 AA; 108669 MW; ECB2E9E92E382853 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 866
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           THERMOPSIN PRECURSOR RELATED PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (TrEMBLrel. 16, C
(TrEMBLrel. 16, I
(TrEMBLrel. 19, I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; AL445063; CAC11313.1;
MEROPS; A05.UPW; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   407:508-513(2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=2303;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-MAR-2001 (
01-MAR-2001 (
01-DEC-2001 (
                                                                                                                                                                                                                                                                                                                                                                                                                                       683 AYR 685
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     818 SYR 820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Nature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 14
29HLQ9
AC 09HLQ9
AC 09HLQ9
DT 01-MAR.
DT 01-DEC
DE THERMOIO
CO Thermol
OX NCBL_TI
RN [1]
RN [1]
RN KEDLIN
RX MEDLIN
RY MEDLIN
RY MEDLIN
RY MENCOPOLOR
DR EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              601H60
                                                                                                                                                                                    493
                                                                                                                                                                                                                                                                                                           581
380
                                                            413
                                                                                                                                                                                                                                                 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
à
                           g
                                                                                    qq
                                                                                                                     ò
                                                                                                                                                  g
                                                                                                                                                                                 δ
                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                         Ω
                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                          ά
                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
379
                                                                                                                                                                                                                                                  269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         645
                                                                                                                                                                                                                                                                                                                                                                                                                   438 KIGDIYAPVVIENYSDALYYSRSGNGSQISPFVISFGQIVSVDPLFGVLNDYLFPVFIGL 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        356 DISAEPWLTFGVQPNPPVPSPKLGWMDEAM------AIDPFNSDRMLYGTGATL 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       451 L---GGFTHADVTAVPS-TIFTSPVFTTGTSVDYAELNP---SIIVRAGSFDPSSQPNDR 503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        700 IVGGGYEGGNYW--------GSPFITPPWNESGYIDSGYDYHPLVYFL 739
---DSGATW 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          498 AIINTSYHIDVENASSFEITYGSQY--DSILASHGLGNTNWMGMILYGTS--NVSITGGL 553
                                      331 GETSYGVDVGWY-GNMADLQSGPSLIYG-LW-----NVSSGIETI-SGTAEPQSYIF
                                                                                                                                                                                                                                                                                                               380 VSNGKQFNNSTASWAPYSGAHFRFKLPKGSYSLA--VLRNLYAPEYINDAGTSPSLNGTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      222 WS-----RDGGATWQAVPGAPTGFIPHKGVFDPVNHVL-----YIATSNTGGPYDGSS
                                                                                                                                                                                                                                                                                                                                                                      270 --GDVWK-----FSVTSGTWTRISPV------PSTDTANDYF--GYSGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                304 T-----IDRQHPNTIMVATQISWWPDTIIFRSTDGGATWT-RIWDWTSYPNRSLRYVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               404 YATNDLTKWDSGGQIHIAPMVKGLEETAVND------LISPPSGAPL---ISALGD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 504 HVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFVWAP------GDPGQPVV---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     552 YAVGFGNSWAASQGVPANA--QIRSDRVNPKTFYALSNGTFYRSTD---GGVTFQPVAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----AVGVMFHAVPGKEGDLW-LAASSGLYHSTN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    847 QDLTVSLPNGTYS----YSVFAPPNNYTAVSGYVTVNGVSEIVNITFPEQLYRVVIRET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomyces.
NCBI_TaxID=1902;
                                                                                                                       162 SQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----GWFPETMKGFVVANLMVWNSSNDLIADNVFY---SL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    641 GGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVT-----GAYR----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-A3(2);
Seeger K.J., Harris D.;
Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Last sequence update)
Last annotation update)
113 PLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       848 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
```

```
41;
                                                                                                                                                                                                                                                                                                                                                              Kinashi H., Hopwood D.A.;

"A set of ordered cosmids and a detailed genetic and physical map for the 8 Mo Streptomyces coelicolor A3(2) chromosome.";

MOL Microbiol. 21:77-96(1996).

EMBL, AL132662; CAB59602.1;

InterPro; IPR000421; FA58_C.

InterPro; IPR001092; HHL4dim.

Pfam: PF00754; F5_F8_type_C; 1.

SMART; SM00231; F358C; 1.

SEQUENCE 848 AA; 89530 MW; 13CEB7C7F241B231 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                297 ---TGAPPVEGGGD-LGPNVIVVDPSTPNLQQKFDQVFAQQESNQF----GGGRXQFLLK 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     404 AITPSNCTDRWAVAQAAPFRRIHVKGGLNLAPNGY-GWASGGYI------ADSKIDG 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           73 ADPINTNKVWAAVGMYTNSWDPNDGA--ILRSSDQG-----ATWQITPLPFKL----- 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     119 ---GGNMPG-----RGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPD 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    283 TRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWD 342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  343 WTSYP--NRSLRYVLDISAEPWLTFGVQPNPPVPSPKL----GWMD-------EAM 385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              349 PGTYNGINAQLGFYTSVSG----LGLNPDDTQINGDITVDAGWFNGNATQNFWRSAENL 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             488 IVRAGSFD---PSSQPNDRHVAFSTDGG-----KNWFQGSEPGGVTTGGTVAASAD-GSR 538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGV 598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              599 TFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSALTGVSSAVNVG 658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        103 QD----GSTWTDAKTVTG----GDGGTDTLDVSGQGRYVRMQGVHRATPWGYSLWEFQVF 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  155 GTTGGTAPGGSCDTANAAQGRPATASSVENA---GTPATAAF--DGDTGTRWSSQASDPQ 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                386 AIDPFNSD------RMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEETAVND 434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               435 LISPPSGAPLISALGDLGGFTH-----ADVTAVPSTIFTSPVFTTGTSVDYAELNPSI 487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            572 LLFTPG-----VYHI-------DQTIEVDRANT---VVLGLGLATIIPDGGV 608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               609 DAMHV-ADVDGVRLAGFLIDAGPVNSDTLLRIGTPG-----GNADHAANPTTMQDVFIR 661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        25 NEGAPGILYVRTDIGGMYRWDAANGRWIP------LLDWVGWNNWGYNGVVSIA 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ------DLNWEAAYAKDFRIE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         240 ASADGOSWQTLKSV-TGATGGKASYDVSGSGRYVRVHGTARA-TGYGYSLWEVAVHTGT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---GGATWQAVPGAPIGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          514 YLDGNEYKVRVPAKRTNARGVSWPANAGGTSLPLSQFYVVKPG--ATAATINAALDQGLN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   170 VGTYIANPIDTIGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRD----
                                                                                                                                                                                                                                                                                                                                Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J., Kinashi H., Hopwood D.A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels 219;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 848;
                                      STRAIN-A3(2);
Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 4.5%; Score 182.5; DB 2; Best Local Similarity 20.1%; Pred. No. 0.021; Matches 163; Conservative 94; Mismatches 336;
                                                                                                                                                                                                                                                   STRAIN=A3(2);
MEDLINE=97000351; Pubmed=8843436;
                                                                                                                                                                                                    SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      226
RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRER RRARRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           φ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
```

us-09-917-376-3.rspt

•

Run

```
Cellulase AE-1. A
Truncated cellulas
Truncated cellulas
Truncated cellulas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Cellulose binding
A mannanase-linker
Pectate lyase-link
Pectate lyase CBD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NK-1 cellulase. B
Novel endoglucanas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Corrected Bacillus
                                                                                                                  2, 2002, 09:15:50 ; Search time 101.03 Seconds (without alignments) 97.848 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Description
                                                                                                                                                                                                                           470
1 VSGGVKVQYKNNDSAPGDNQ......RASFGSVNPATPTADTYLQX
                                                                                                                                                                                                                                                                                                                                                                                            747574
GenCore version 4.5
Copyright (c)·1993 - 2000 Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                                          hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                 747574 segs, 111073796 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Maximum Match 100%
Listing first 45 summaries
                                                                               OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAR15625
AAY13494
AAY13493
AAY13492
AAR13227
AAR13227
AAR13227
AAR95080
AAX54123
AAX54123
                                                                                                                                                                                                                                                                                     BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A_Geneseq_032802:*
                                                                                                                                                                                                                                                                                                                                                                                                                                   seq length: 0 seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Post-processing: Minimum Match 08
                                                                                                                                                                                                        US-09-917-376-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               图
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       616
1751
1426
499
700
551
167
476
493
                                                                                                                        July
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       441.6
441.6
339.74
335.1
335.1
355.1
355.1
355.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   66...
1111...
1111...
1111...
1111...
1111...
1111...
1111...
1111...
1111...
                                                                                                                                                                                                                                                                                                                                                                                            οŧ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               231.5
1994.5
191.5
191.5
186
172
165.5
165.1
165
165
                                                                                                                                                                                                                           Perfect score:
                                                                                                                                                                                                                                                                                       Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                          Total number
                                                                                                                                                                                                                                                                                                                                                                                                                                 Minimum DB
Maximum DB
                                                                                                                                                                                                                                                 Sequence:
                                                                                                                                                                                                                                                                                                                                                 Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Database
                                                                                                                          on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Result
No.
```

1. to 1

us-09-917-376-4.rag

```
Η;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The invention relates to a recombinant cellulase active protein free of proteinases of native thermophilic and alkalinophilic origin, comprising the truncated sequences Cel B5, Cel B4/5, Cel E1, Cel E1/2, Cel 1/2/3, Cel 6 or Cel E3/B5, or a stability region from one of the defined full-length sequences, or functional equivalents. Cel B5 extends from amino acid A1011 to P1424 or K1425 or N1426, and Cel B4/5 extends from amino acid A1012 in the sequence shown in AAV13492; Cel E1 extends from amino acid V39 to D481, Cel E1/2 extends from Y39 to G812, Cel E6 extends from y39 to G812, Cel E6 extends from wino acid V1233 to K1751 and the stability region extends from amino acid V1233 to the sequence shown in AAX13493; Cel E3/B5 is shown in AAX13494. The new enzymes are useful in laundry detergent compositions to prevent or
                                                                                                                                                                                                                                                                                      E1/2;
                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                   61
gene (obtd. from Aeromonas) into pUC18. The protein has amol. wt. of 81,000 (SDS-PAGE) and an optimum pH near to 5 when carboxymethylcellulose is the substrate. The N-terminal sequence: GIHADThas been confirmed by Edman degradation. The gene can be used to produce recombinant enzyme which is used for the effective utilisation of biomass resources and the mfr. of pharmaceuticals and foodstuffs, and also for the detergent and deinking of waste paper.
                                                                                                                                                                                                                                                                 2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Cellulase; proteinase; truncated; Cel B5; Cel B4/5; Cel E1; Cel Cel 1/2/3; Cel 6; Cel E3/B5; laundry detergent; stain-removing; cotton-containing fabric; stonewashing.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New truncated cellulase proteins, useful in detergents and for producing 'stonewashed' denim
                                                                                                                                                                                                Length 782;
                                                                                                                                                                                            49.3%; Score 231.5; DB 12; Length 48.3%; Pred. No. 5e-18; Live 16; Mismatches 28; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Farrington GK;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Daniels RM,
                                                                                                                                                                                                                                                                                                                                                         62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Claim 7; Page 42-43; 65pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY13494 standard; Protein; 616 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PL, Daniel Williams DP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (CLRN ) CLARIANT FINANCE BVI LTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Truncated cellulase Cel E3/B5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98EP-0810919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         97US-0932571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                         Best Local Similaring
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bergguist
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Morgan H,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1999-315403/27
                                                                                                                                            782 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; AAX55660.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15-SEP-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19-SEP-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30-JUL-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EP921188-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Anderson P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gibbs MD,
                                                                                                                                              Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY13494;
                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY13494
                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            88888888888
                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                     ò
```

```
ä
                   useful
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           proteinases of native thermobilic and alkalinophilic origin, comprising the truncated sequences Cel B5, Cel B4/5, Cel E1, Cel E1/2, Cel 1/2/3, Cel 6 or Cel E3/9E, or a stability region from one of the defined full-length sequences, or functional equivalents. Cel B5 extends from amino acid A1011 to P1424 or K1425 or N1426, and Cel B4/5 extends from amino acid K635 to N1426 in the sequence shown in AAY1349; Cel E1 extends from amino acid X195 to D481. Cel E1/2 extends from X39 to G635. Cel E1/2/3 extends from Y39 to G812. Cel E6 extends from A79 to G635. Lel E1/2/3 extends from X39 to G812. Cel E6 extends from amino acid V1233 to the sequence shown in AAX13493; Cel E6 extends from amino acid V1233 to the sequence shown in AAX13493; Cel E3/55 is shown in AAX13494. The new enzymes are useful in laundry detergent compositions to prevent or remove staining, backstaining or graying, for use on cellulosic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The invention relates to a recombinant cellulase active protein free of
remove staining, backstaining or graying, for use on cellulosic materials including cotton-containing fabrics. They are especially use for preventing redeposition of colorant during stonewashing, and for processing of textiles where cellulose breakdown is required. The new truncated enzymes show reduced redeposition of dye compared to using non-truncated cellulase compositions.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Cellulase; proteinase; truncated; Cel B5; Cel B4/5; Cel E1; Cel E1/2; Cel 1/2/3; Cel 6; Cel E3/B5; laundry detergent; stain-removing; cotton-containing fabric; stonewashing.
                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                              1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                    ;
                                                                                                                                                                                      Length 616;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  proteins, useful in detergents and denim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Anderson P, Bergquist PL, Daniels RM, Farrington GK;
Gibbs MD, Morqan H, Williams DP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Truncated cellulases comprising amino acid sequence.
                                                                                                                                                                                    DB 20;
                                                                                                                                                                                    ; Score 195.5; DB 20; Pred. No. 5.4e-14; 15; Mismatches 34.
                                                                                                                                                                                                                                                                                                                                          61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                             waqigasnvtfnfvklssgvsgadyyle 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY13493 standard; Protein; 1751 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 7; Page 37-41; 65pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (CLRN ) CLARIANT FINANCE BVI LTD
                                                                                                                                                                                      41.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98EP-0810919.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 97US-0932571.
                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      New truncated cellulase producing 'stonewashed'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1999-315403/27
                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 38; Conserv
                                                                                                                                616 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-PSDB; AAX55662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15-SEP-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19-SEP-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 30-JUL-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EP921188-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       09-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY13493;
                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY13493
 8888888
                                                                                                                                                                                                                                                                  ð.
                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                             g
```

 $\sim$ 

```
The invention relates to a recombinant cellulase active protein free of proteinases of native thermophilic and alkalinophilic origin, comprising the truncated sequences Cel B5, Cel B4/5, Cel E1/2, Cel E1/2/3, Cel E1/2, Cel E1/
                                                                                                                                                                                                                                                                                                Η;
materials including cotton-containing fabrics. They are especially useful for preventing redeposition of colorant during stonewashing, and for processing of textiles where cellulose breakdown is required. The new truncated enzymes show reduced redeposition of dye compared to using non-truncated cellulase compositions.
                                                                                                                                                                                                                                                                                                                                                                            Cel B4/5; Cel E1; Cel E1/2;
                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                              GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Cellulase; proteinase; truncated; Cel B5; Cel B4/5; Cel E1; Cel Cel 1/2/3; Cel 6; Cel E3/B5; laundry detergent; stain-removing; cotton-containing fabric; stonewashing.
                                                                                                                                                                                                                                               DB 20; Length 1751;
                                                                                                                                                                                                                                                                                                ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            truncated cellulase proteins, useful in detergents and for lucing 'stonewashed' denim
                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Farrington GK;
                                                                                                                                                                                                                                       Query Match 41.4%; Score 194.5; DB 20; Best Local Similarity 44.7%; Pred. No. 2.6e-13; Matches 38; Conservative 14; Mismatches 32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Truncated cellulases Cel B4/5 and Cel B5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Daniels F
ams DP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 7; Page 34-37; 65pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      737 igasnvtfnfvklssgvsgadyyle 761
                                                                                                                                                                                                                                                                                                                                                                                                                                              64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY13492 standard; Protein; 1426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (CLRN ) CLARIANT FINANCE BVI LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PL, Dan
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            97US-0932571.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98EP-0810919.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bergquist
Morgan H, W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1999-315403/27.
N-PSDB; AAX55661.
                                                                                                                                                             1751 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15-SEP-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19-SEP-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         30-JUL-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EP921188-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       09-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Anderson P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         producing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY13492;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gibbs MD,
                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                              4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nev
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY13492
     888888888
                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                             d
                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

```
ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ä
             The new
                                                                                                                                                                         for preventing redeposition of colorant during stonewashing, and for processing of textiles where cellulose breakdown is required. The new truncated enzymes show reduced redeposition of dye compared to using non-truncated cellulase compositions.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequences (AAQ49818-19) consist of two synthetic fragments which are used to induce a mutation within the cellulase NK-1 gene. The gene is shown in sequence (AAQ49820). The modified NK-1 gene shows a change in pH dependence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA 62
                                                                                                                                                         4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Changing the pH-dependence of cellulase enzymatic activity - by changing base sequence of cellulose-producing gene of Bacillus microbe to base sequence coding asparagine and serine residues at specified aminoacid sites.
                                                                                                         Length 1426;
                                                                                                                                 ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 499;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                       20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           39.6%; Score 186; DB 14; 40.7%; Pred. No. 5.2e-13;
                                                                                                         Score 191.5; DB 2
Pred. No. 4.4e-13;
                                                                                                                                15; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                             Cellulase; pH dependence; mutation.
                                                                                                                                                                                                                                     496
                                                                                                                                                                                                                                                                                                       AAR42122 standard; Protein; 499 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                            64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Disclosure; Fig 1; 9pp; Japanese.
                                                                                                                                                                                                                                 472 igasnvtfnfvklssgvsgadyyle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19;
                                                                                                       40.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    92JP-0075883.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92JP-0075883
                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                      Query Match
Best Local Similarity 43.5 Matches 37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1993-330585/42.
N-PSDB; AAQ49820.
                                                                  AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
Matches 35; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AA;
                                                                1426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (BEPP/) BEPPU T.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  499
                                                                                                                                                                                                                                                                                                                                                                                   NK-1 cellulase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JP05236969-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           17-SEP-1993.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-FEB-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-FEB-1992;
                                                                                                                                                                                                                                                                                                                                                         27-APR-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacillus N4
                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                 AAR42122;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                        AAR42122
                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                 8×66666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                          à
                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

3A.

AAR13227;

AAR13227

RESULT

g

Peptide Protein

```
2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The present sequence is the corrected version of the incorrect Bacillus lautus (NCIMB 40250) endoglucanase Endo 3A described in W09110732. Endo 3 can be used in novel method of forming localised colour density variation on the surface of a dyed cellulosic fabric. The method comprises agitating the fabric in an aqueous medium (PH 6.5 to 9.0) containing a family 5 cellulose.

1.4 cellobioside, or a family 7 cellulase, and a mechanical abrading agent or cellulose having abrading activity. Each cellulase displays 30 % or more of its maximum activity at ph 7. The process is useful to provide a stone washed look to blue jeans
                                                                                                                                                                                                                                                                                                                                                                                       Endoglucanase; Endo 3A; formation; localised; variation; colour density; surface; dye; fabric; family 5; cellulose; hydrolysation; p-nitrophenyl-beta-1,4-cellobioside; stone wash; blue jeans; back staining.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Localised variation of colour density in the surface of a dyed cellulosic fabric - uses cellulase compsn. able to hydrolyse
                                                                                                                                                                                                                                                                                                                   Corrected Bacillus lautus (NCIMB 40250) endoglucanase Endo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                35.2%; Score 165.5; DB 18; Length 37.9%; Pred. No. 1.4e-10; Live 22; Mismatches 29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Toft AH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Disclosure; Pages 15-17; 23pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           p-nitrophenyl -beta-1,4-cellobioside
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | :| |: | :| | 461 aqigasnvsaaf--anftgsntdtyve 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                      AAW18790 standard; protein; 551 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAR95080 standard; peptide; 167 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Schulein M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          96WO-DK00364.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       95DK-0000993
                                                                                                                                                                                                                             18-NOV-1997 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 37.9%
Matches 33; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    30-OCT-1996 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (NOVO ) NOVO-NORDISK AS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                without back staining
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Fich M, Onishi M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1997-192888/17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      551 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacillus lautus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO9709410-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      03-SEP-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       08-SEP-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13-MAR-1997.
                                                                                                                                            AAW18790;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR95080;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAR95080
ID AAR9
XX
AC AAR9
XX
DT 30-00
             AAW18790
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                                                                                                                                            X D D X D D X D X D X D X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X D X D X D X X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Э;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 VSGGVKVQYKNND-SAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNC 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The enzyme is produced by a strain of Bacillus spp. NCIMB 40250 and exhibits an endoglucanase activity of at least 10 (pref. at least 25) carboxymethyl cellulose (CMC) endoase units per mg total protein under alkaline conditions. It is especially useful as a cellulolytic agent and has been found to be more stable during washing (60 mins. at 40 deg.) in the presence of conventional detergents than a commercial cellulase preparation. It may also show increased storage stability in lig. detergents contg. proteases. The sequence was deduced from the DNA (AAQ13001). See also AAR13228 and AAR13229.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Enzyme exhibiting cellulase activity from Bacillus sp. - is an endo-glucanase, esp. useful for harshness redn. of cotton-contg.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ,
4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 36.6%; Score 172; DB 12; Length 700; 44.9%; Pred. No. 3.3e-11; Live 18; Mismatches 27; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Location/Qualifiers
1.31
/label= signal sequence
32.700
/label= mature endoglucanase
31.32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ||| :| ||: |||: | | : ||| : ||| : ||| : ||| : ||| : ||| : ||| | : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : || : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : || : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : || : ||| : ||| : ||| : ||| : ||| : ||| : ||| : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DWAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hansen C;
410 gigcgnlthkfvtlhkpkggadtyle 435
                                                                                                                                                                                                                    AAR13227 standard; Protein; 700 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Claim 1; Page 80; 96pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Cellulase activity; detergent.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                91WO-DK00013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 90DK-0000164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Jorgensen PL, Schulein M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 36.6%
Best Local Similarity 44.9%
Matches 40; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacillus spp. NCIMB 40250.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (NOVO ) NOVO NORDISK A/S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1991-238020/32.
N-PSDB; AAQ13001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Novel endoglucanase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       700 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Cleavage-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                18-JAN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19-JAN-1990;
                                                                                                                                                                                                                                                                                                                                                                                                    14-OCT-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO9110732-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25-JUL-1991
```

fabrics

ñ

Gaps

3;

Seguence

9 609

ò g ö RESULT

DB 18; Length 551;

S

```
ä
                                                                                                                                                                                                                                                                                                                                                                   A modified cellulose binding domain (CBD) or fraction of it, linked to a hapten moiety via one or more Cys or Lys residues and complexes comprising the biotinylated CBD and a biotin-binding molecule selected from modified or unmodified avidin or streptividin or an anti-biotin antibody, can be used in affinity chromatography, cell separation, cell, protein and enzyme immobilisation, selective removal of bological materials, multienzyme reactors, signal immunoassays and drug (e.g. pesticide) delivery. The CBD molecule can be modified with the hapten without affecting its high affinity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                  Cellulose binding domain; CBD; hapten; moiety; biotin; avidin; streptavidin; affinity chromatography; cell separation, cell immobilisation; protein immobilisation; enzyme immobilisation; multienzyme reactors; signal immunoassays; drug delivery; pesticide;
                                                                                                                                                                                                                                                                                                         Modified cellulose-binding domain mols. - having attached hapten, partic. biotin, to provide an affinity system for, e.g. sepns., assays, reactors, delivery etc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 167;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A mannanase-linker-cellulose binding domain fusion protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                  Morag E, Shoham Y, Wilchek M;
                                                                                                                     /note= "Unidentified amino acid."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                35.1%; Score 165; DB 17; 36.1%; Pred. No. 3.6e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                        (TECR ) TECHNION RES & DEV FOUND LTD.
(UYRA-) UNIV RAMOT APPLIED RES & IND DEV LTD.
(YEDA ) YEDA RES & DEV CO LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    23; Mismatches
                                                                                                                                                                                                                                                                                                                                                  Claim 12; Page 34-35; 53pp; English.
                                                                                                 Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY54123 standard; Protein; 476 AA.
                                                                                                                                                                                  95WO-US13813.
                                                                                                                                                                                                      94IL-0111415.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 35.1%
Best Local Similarity 36.1%
Matches 35; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                               Clostridium thermocellum.
Cellulose binding domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              cellulose, or chitin
                                                                                                                                                                                                                                                                  Bayer EA, Lamed R,
                                                                                                                                                                                                                                                                                     WPI; 1996-239453/24.
                                                                                                            Misc-difference 139
                                                            cellulose; chitin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 167 AA;
                                                                                                                                                                                 26-OCT-1995;
                                                                                                                                         WO9613524-A1
                                                                                                                                                                                                     27-OCT-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             27-MAR-2000
                                                                                                                                                             09-MAY-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY54123;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY54123
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEXTXEX
```

```
Mannanase; mannan endo-1,4-beta-mannosidase; beta-mannanase; endo-1,4-mannanase; Balatchomnan; aladoct-mannanase; Balatchomnan; 1,4-beta-b-mannosidic linkage; mannan; galatchomannan; glucomannan; galactoglucomannan; cellulosic fibre; synthetic fibre; yarn; fabric; parlating paste; plant material degradation; recycled waste paper; paper making pulp; guar; locust bean gum; thickener; viscosity; mannan-containing food; coffee extract; cleaning composition; machine washing; hard-surface cleaner; dishwashing; oral; dental; contact lens; body-care composition; fabric softener; oil well drilling; subterranean formation fracture; cellulose binding domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           domain fusion protein. Mannanase (also known as mannan endo-1,4-beta-mannosidase, beta-mannanase, or endo-1,4-mannanase) hydrolyses galactomannans. Specifically, mannanases hydrolyses labortomannans. Specifically, mannanases hydrolyse and galactomannans. The mannanase protein, or preparations and galactoglucomannans. The mannanase protein, or preparations containing it, are used to improve properties of cellulosic or synthetic fibres, yarn or (non)wown fabrics (removal of mannan-based sizes or printing pastes). They are also used to degrade or modify plant materials (particularly recycled waste paper, paper making pulps, or material containing quar or locust bean gums (thickeners), or to reduce viscosity of mannan-containing foods or feeds). The mannanases are also used to process coffee extracts (to inhibit gel formation); in cleaning compositions (for machine washing of fabrics, as
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hard-surface cleaners, for hand or machine dishwashing, also in oral, dental, contact lens or body-care compositions) where they remove mannan-containing soils and prevent binding of some soils to cellulosics; and in fabric softeners. They can also be used in oil well drilling to fracture subterranean formations.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The present sequence represents a mannanase-linker-cellulose binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New mannanases for treatment of textiles, plant material and coffee
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bjornvad ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 476;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Andersen LN,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 165; DB 21;
Pred. No. 1.3e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Example 4; Page 211-212; 242pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Schnorr K,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 extract, and in cleaning compositions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Kauppinen MS, Schuelein M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          35.1%;
36.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98US-0105970.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98DK-0001725.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99WO-DK00314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99DK-0000308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99DK-0000309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0123543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0123623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0123642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98DK-0001341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99DK-0000307
                                                                                                                                                                                                                                                                                                                                                                                                                                          Clostridium thermocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (NOVO ) NOVO-NORDISK AS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2000-105891/09.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-PSDB; AAZ45336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO9964619-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-OCT-1998;
28-OCT-1998;
23-DEC-1998;
05-MAR-1999;
05-MAR-1999;
05-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                              Bacillus sp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10-MAR-1999;
11-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10-JUN-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                20-0CT-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   20-0CT-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      09-MAR-1999
10-MAR-1999
                                                                                                                                                                                                                                                                                                                                                                           Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
```

셤 õ

Ā

(first entry)

```
AAY43218 standard; Protein; 493
                                                                                                                                                        13-JAN-2000
                                                                                                                         AAY43218;
                                                             RESULT 11
                                                                             AAY43218
                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           plasmid pMB914 was constructed using pectate lyase gene from
Bacillus licheniformis, ATCC 14580 and cellulose binding domain (CBD)
gene from clostridium thermocellum YS fused with a linker molecule.
Bacillus subtilis was transformed with plasmid pMB914 for expression of
the fusion protein. Pectate lyase can be used in detergent compositions,
for cleaning hard surfaces, for machine treatment of fabrics, for
improving the properties of cellulosic fibres, yarn, woven or non-woven
fabric, for the degradation of plant material e.g. recycled waste paper,
mechanical paper-making pulps or fibres subjected to retting process, for
preparing animal feed and for processing wine or juice.
                                                                                                                                                                                                                                                                                                                        Pectate lyase-linker-CBD fusion protein; ATCC 14580; plasmid pMB914; cellulose binding domain; CBD; linker; transformed; detergent; cellulosic fibre; yarn; degradation; recycled waste paper; animal feed; paper-making pulp; retting process; processing; wine; juice.
                                              1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The present sequence is a Pectate lyase-linker-CBD fusion protein.
                              1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Lange NEK, Schnorr K, Schuelein M;
 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 493;
29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  35.1%; Score 165; DB 20; 36.1%; Pred. No. 1.4e-10; tive 23; Mismatches 25
                                                                                                                      373 haaiigsngsyngitsnvkgtfvkmssstnnadtyle 409
                                                                                          WAAM------GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WAAM------GCGNIRASFGSVNPATPTADTYLQ 88
23; Mismatches
                                                                                                                                                                                                                                                                                            Pectate lyase-linker-CBD fusion protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Example 4; Page 85-86; 93pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                     Chimeric - Bacillus licheniformis.
Chimeric - Clostridium thermocellum.
                                                                                                                                                                                                     Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New isolated pectate lyase enzymes
                                                                                                                                                                                                  AAY28850 standard; Protein; 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98WO-DK00514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          97DK-0001344
98US-0073684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Andersen LN, Bjornvad ME,
                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (NOVO ) NOVO-NORDISK AS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1999-610578/52.
N-PSDB; AAX90978.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity
nes 35; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      493 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   W09927083-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            36-MAY-1998;
                                                                                                                                                                                                                                                               17-JAN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 03-JUN-1999
35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                AAY28850;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61
Matches
                                                                                                                                                                                   AAY28850
```

a ò

ò

```
'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This sequence is a fusion protein comprising a Bacillus species pectate lyase of the invention, fused via a linker to the Clostridium thermocollum Cipb protein cellulose binding domain (CBD). The pectate lyase enzymes are obtained from Bacillus agaradhaerens (BA), Bacillus licheniformis (BL), Bacillus halodurans (BH) and other Bacillus species. The pectate lyase enzymes can be used for degrading pectin, pectate and polygalacturonides. They can be used in detergent compositions, for cleaning hard surfaces, for machine treatment of fabrics, for improving the properties of cellulosic fibres, yarn, woven or non-woven fabric, for the degradation of plant material (e.g. recycled waste paper, mechanical paper-making pulps or fibres subjected to a retting process, for preparing animal feed or for processing wine or juice. DNA encoding the enzymes can also be used for the production of transgenic plants.
                                                          Pectate lyase; polysaccharide lyase; enzyme; pectin degradation; polygalacturonide; detergent composition; hard surface treatment; cellulosic fibre; plant material degradation; recycled waste paper; mechanical paper-making pulp; wine processing; cellulose binding domain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 493;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              , Lange NEK, Bjornvad ME,
Schnorr K, Kongsbaek L;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        35.1%; Score 165; DB 20;
36.1%; Pred. No. 1.4e-10;
tive 23; Mismatches 29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 WAAM------GCGNIRASFGSVNPATPTADTYLQ 88
Pectate lyase CBD fusion protein sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 26; Page 106-108; 113pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New isolated pectate lyase enzymes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          97DK-0001343.
97DK-0001344.
98US-0073684.
98US-0184217.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98WO-DK00515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Schuelein M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                    Clostridium thermocellum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Kauppinen MS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NOVO ) NOVO-NORDISK AS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1999-610579/52.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
Matches 35; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   493 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-PSDB; AAZ31562
                                                                                                                                                                                                                                                                                                                                                                          WO9927084-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 24-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            02-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Andersen LN,
                                                                                                                                                                                                                                                                                                                                                                                                                                        03-JUN-1999
                                                                                                                                                                                                                                                                                     Bacillus sp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               06-MAY-1998
                                                                                                                                                                                                                                                                                                               Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Glad SOS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
```

531 AA;

Sequence

S

셤

```
This protein sequence represents a scaffoldin based on the CipA protein of Clostridium thermocellum. The scaffoldin protein is used in a novel composition that comprises at least 2 enzymes non-covalently bound to a peptide backbone (i.e. present sequence). The scaffoldin comprises a number of internal repeating units and at least one cellulose binding domain (CBD). The CBD may be altered to modify its affinity for cellulose, which may be desirable where cellulose binding would be disadvantageous. The enzyme is bound to the scaffoldin by a dockerin cellulose, which binds to a repeating element of the capton of the enzyme, which binds to a repeating element of the scaffolding. The dockerin is preferably Cels (AAM1523) or Celb scaffolding. The dockerin is preferably Cels (AAM1523) or Celb (AAM15236). The composition can be used in reducing allergenicity, producing synergistic effects, and facilitating selective modification of substrate. By taking advantage of the cellulose binding domain of the scaffolding domain, it is possible to recover enzymes, or to quantify the amount of an enzyme delivery system which could be used in the food industry, in food processing, animal feed, textiles, bioconversion, pulp and paper production, plant protection and pest control, as a wood control, as a wood control or control, as a wood control or control or control or control or control or control or control c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Composition with enzymes non-covalently bound to a peptide backbone - used as a multi-enzyme delivery system, e.g. in food processing,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Dockerin; CelD; CelS; CipA; scaffoldin; cellulose binding domain; chromatographic separation; soluble substrate modification; CBD; multi-enzyme delivery system; animal feed; paper production; plant protection; pest control.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Xia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              154..306
/label= internal_repeat_element_2
239..51
/label= cellulose_binding
/note= "only 60% of the CBD is present"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ward M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1..153
/label= internal_repeat_element_1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Scaffoldin protein from Clostridium thermocellum.
Liu C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
                                                                                                                                                                                                                                                                AAW15238 standard; protein; 531 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Clarkson KA, Fowler T,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 10; Fig 6; 29pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        95US-0559968.
95US-0005701.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    96WO-US16485.
                                                                                                                                                                                                                                                                                                                                                                                                                                            28-JAN-1998 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   textiles and pest control
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Clostridium thermocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GEMV ) GENENCOR INT INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1997-245106/22.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO9714789-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16-OCT-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    17-NOV-1995;
17-OCT-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24-APR-1997
                                                                                                                                                                                                                                                                                                                                                        AAW15238;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bott RR,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Domain
                                                                                                                                                                            RESULT 12
                                                                                                                                                                                                                      AAW15238
```

```
Multimeric protein, especially enzymatic, complexes are held together by protein protein interactions between domains designated dockerins and cohesins, which are found on the catalytic and scaffold subunits respectively. An example of such a complex is the cellulose degrading protein complex from Clostridium thermocellum, known as the cellulosome. This complex comprises around 15 protein fainluding endoglucanses, cellobiohydratases, hemicellulases, e.g. xylanases or lichinases, which interact with a central "scaffold" protein designated the cellulosome interact with the cip subunit via conserved 23 amino acid dockerin domains. Cip has been shown to contain 9 copies of a cohesin domain. The invention relates to the isolation of proteins blading to a novel dockerin type domain found in the C-terminal portion of Cip. The new domain is designated a type II dockerin domain (as compared to the type II dockerin domains.) The store on the catalytic subunits of the cellulosome). The type II dockerin domains.

The sequence presented here is the cellulosome integrating protein to sequence similarity to the type II dockerins but is unable to bind type I cohesin domains.

The sequence presented here is the cellulosome integrating protein to bind type I cohesin domains.

The sequence presented here is the cellulosome integrating protein bund at the C-terminal dockerin type II domain (location not given in specification). The novel type II dockerin and cohesin domains can be
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Multimer; enzyme; complex; protein-protein interaction; dockerin domain; cohesin domain; catalytic subunit; scaffold subunit; SdbA; synergistic; cellulosome integrating protein; scaffoldin dockerin binding protein.
                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Cellulase proteins with cohesin or dockerin type II domains - useful for potentiating the activity of multiprotein enzyme complexes
                                                                                    1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                          10;
Length 531;
                                            Indels
35.1%; Score 165; DB 18;
36.1%; Pred. No. 1.5e-10;
iive 23; Mismatches 29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     C. thermocellum cellulosome integrating protein.
                                                                                                                                                                                              ||: :| :| |||||: 371 haailgsngsyngitsnvkgtfvkmssstnnadtyle 407
                                                                                                                                                                     61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                          AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Disclosure; Page 47; 60pp; French.
                                                                                                                                                                                                                                                                                                                     AAW43108 standard; Protein; 1853
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         96FR-0005854.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   96FR-0005854.
                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
  Query Match 35.19
Best Local Similarity 36.13
Matches 35; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Clostridium thermocellum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (INSP ) INST PASTEUR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Leibovitz
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI: 1998-011569/02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-PSDB; AAT86625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FR2748479-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10-MAY-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10-MAY-1996;
                                                                                                                                                                                                                                                                                                                                                                                                           16-0CT-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 14-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Beguin P,
                                                                                                                                                                                                                                                                                                                                                                 AAW43108;
                                                                                                                                                                                                                                                                                RESULT 13
                                                                                                                                                                                                                                                                                                      AAW43108
                                                                                      ð
                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                δλ
```

ထ

Page

1;

Gaps

ij

Indels

Pred. No. 5.9e-09;

31.8%; Preu. .... +ive 21; Mismatches

S S X S

ò

ò

```
4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                           64 MGCGNIRASFGSVNPATPTADTYLQ 88
              Conservative
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            30-AUG-2001
                                                                                                                                                                                                                                                     AAG63963;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                          15
                Matches
                                                                                                                                                                                                       AAG63963
                                                                                                                                                                                                                                                         g
                                               δ
                                                                                                            ò
                                                                                                                           ;
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          specification describes a xyloglucanase enzyme belonging to family 44 of at least 196 at ph 5-8. The enzyme exhibits high performance in detergent compositions and prevents binding of certain soils to the xyloglucan left in the cellulosic material. It is stable at pH 5-10 at room temperature and has a half life of more than 50 days when incubated in a full formulated liquid detergent at 30 plus degrees celsius. The enzyme is used in detergent compositions, textile industry for improving the properties of cellulosic fibers, yann, and woven or non-woven fabrics, preferably in textile scouring process, and in cellulose fiber propersing industry for ratting of fibers e.g. hemp, jute, flax, and
                                                                                                                                                                             The present sequence represents a xyloglucanase of the invention. The
                                                                                                                                                           1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
used in complexes, especially enzyme complexes, to potentiate their catalytic actions in a synergistic manner.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New xyloglucanase enzyme belonging to glycosyl hydrolases family, useful for detergent compositions, and textile or cellulose fiber processing industries
                                                                                          Length 1853;
                                                                             35.1%; Score 165; DB 19; Lengum 2.36.1%; Pred. No. 7e-10; Wismatches 29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                            xyloglucanase; family 44; glycosyl hydrolase; detergent; cellulosic fiber; textile, scouring.
                                                                                                                                                                                                                                          61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                Amino acid sequence of xyloglucanase enzyme.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Schuelein M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Example 3; Page 81-85; 97pp; English.
                                                                                                                                                                                                                                                                                                                                   AAG63962 standard; Protein; 1352 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          21-FEB-2001; 2001WO-DK00116.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24-FEB-2000; 2000DK-0000291
                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                           35; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Jorgensen PL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Paenibacillus polymyxa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (NOVO ) NOVOZYMES AS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2001-522819/57.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1352 AA;
                                              1853 AA;
                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAH75059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           40200162903-A1.
                                                                                                                                                                                                                                                                                                                                                                                               29-0CT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          30-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Schnorr K,
                                                                                                                                                                                                                                                                                                                                                                 AAG63962;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Seguence
                                               Sequence
                                                                                              Query Match
                                                                                                            Best Local
                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                    AAG63962
```

Length 1352;

DB 22;

33.1%; Score 155.5;

Query Match

```
ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       specification describes a xyloglucanase enzyme belonging to family 44 of glycosyl hydrolases and exhibiting a relative xyloglucanase activity of at least 30% at pH 5-8. The enzyme exhibits high performance in detergent compositions and prevents binding of certain soils to the xyloglucan left in the cellulosic material. It is stable at pH 5-10 at noom temperature and has a half life of more than 50 days when incubated in a full formulated liquid detergent at 30 plus degrees celsius. The enzyme is used in detergent compositions, textile industry for improving the properties of cellulosic fibers, yann, and woven or non-woven fabrics, preferably in textile scouring process, and in cellulose fiber processing industry for ratting of fibers e.g. hemp, jute, flax, and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The present sequence represents a xyloglucanase of the invention. The
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               New xyloglucanase enzyme belonging to glycosyl hydrolases family, useful for detergent compositions, and textile or cellulose fiber processing industries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 1350;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    xyloglucanase; family 44; glycosyl hydrolase; detergent; cellulosic fiber; textile scouring.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 32.0%; Score 150.5; DB 22; 31.8%; Pred. No. 2.2e-08; Live 21; Mismatches 36; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /note= "Glu encoded by GAACCG"
                                                                                                                                                                                                                                                                                                                                                                                                        Amino acid sequence of xyloglucanase enzyme.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Schuelein M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Disclosure; Page 86-89; 97pp; English.
                                                                                                                                                                                                    AAG63963 standard; Protein; 1350 AA.
1259 lscskingklvkmekaatgadyyle 1283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               24-FEB-2000; 2000DK-0000291.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21-FEB-2001; 2001WO-DK00116.
                                                                                                                                                                                                                                                                                                                                        29-OCT-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Schnorr K, Jorgensen PL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 31.8%
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Paenibacillus polymyxa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (NOVO ) NOVOZYMES AS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference 1347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2001-522819/57.
N-PSDB; AAH75060.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1350 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200162903-A1
```

64 MGCGNIRASFGSVNPATPTADTYLQ 88 : | : | | | ||: 1258 lscsklngklvkmdkaatgadyyle 1282

qq δλ

Search completed: July 2, 2002, 09:15:51 Job time: 239 sec

THIS PAGE BLANK (USPTO)

```
July 2, 2002, 09:12:54; Search time 39.87 Seconds (without alignments) 54.524 Million cell updates/sec
                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                      US-09-917-376-4
470
1 VSGCVKVQYKNNDSAPGDNQ.....RASFGSVNPATPTADTYLQX
                                                                                                                                                                                                                                                                                                                                                                                                                                    231628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/pcTUS_COMB.pep:*
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                 Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                  231628 segs, 24425594 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                     OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                 BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                         Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Database :
                                                                                                                                                                                                                                                                                                                                                                                      Searched:
                                                                                                                                   Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Description                   | Sequence 47, Appl | 44,               | Sequence 43, Appl | 2, 1             | Sequence 1, Appli | ف<br>ف           | 10,              | 12,               | Sequence 7, Appli | Sequence 2, Appli | Sequence 2, Appli | 7               | 7,              | Sequence 2, Appli | Sequence 2, Appli | Sequence 7, Appli | Sequence 7, Appli |                  | Sequence 29, Appl |                   | Seguence 5, Appli | Sequence 36, Appl | 3, 4            | Sequence 2, Appli | Sequence 38, Appl | 33,              | Sequence 32, Appl |
|-------------------------------|-------------------|-------------------|-------------------|------------------|-------------------|------------------|------------------|-------------------|-------------------|-------------------|-------------------|-----------------|-----------------|-------------------|-------------------|-------------------|-------------------|------------------|-------------------|-------------------|-------------------|-------------------|-----------------|-------------------|-------------------|------------------|-------------------|
| ID                            | US-09-136-574A-47 | US-09-136-574A-44 | US-09-136-574A-43 | US-07-862-588B-2 | US-09-033-537A-1  | PCT-US95-13813-9 | US-09-198-956-10 | US-09-198-955A-12 | US-07-862-588B-7  | US-08-048-164A-2  | US-08-460-462-2   | US-08-460-457-2 | US-08-460-458-2 | US-08-460-455-2   | US-08-330-394A-2  | US-09-006-636-7   | US-09-006-632-7   | US-09-277-716-22 | US-08-330-394A-29 | US-08-330-394A-22 | US-09-118-319-5   | US-08-353-400-36  | US-09-341-587-3 | US-09-110-517-2   | US-09-423-439-38  | US-08-353-400-33 | US-09-423-439-32  |
| OB                            | . 4               | 4                 | 4                 | ~                | ď                 | വ                | 4                | 4                 | ~                 | _                 | -                 | -               | -               | ~                 | 7                 | ო                 | 4                 | 4                | 7                 | ~                 | m                 | -                 | 4               | 4                 | 4                 | ٦                | 4                 |
| %<br>Query<br>Match Length DB | 616               | 1751              | 1426              | 700              | 551               | 167              | 493              | 493               | 531               | 162               | 162               | 162             | 162             | 162               | 162               | 163               | 163               | 382              | 154               | 156               | 428               | 464               | 1785            | 1581              | 288               | 445              | 673               |
| %<br>Query<br>Match           | 41.6              | 41.4              | 40.7              | 36.6             | 35.2              | 35.1             | 35.1             | 35.1              | 25.6              | 24.3              | 24.3              | 24.3            | 24.3            | 24.3              | 24.3              | 24.3              | 24.3              | 24.0             | 23.8              | 23.8              | 13.6              | 13.6              | 13.5            | 13.4              | 13.3              | 13.3             | 13.3              |
| Score                         | 195.5             | 194.5             | 191.5             | 172              | 165.5             | 165              | 165              | 165               | 120.5             | 114               | 114               | 114             | 114             | 114               | 114               | 114               | 114               | 113              | 112               | 112               | 64                | 64                | 63.5            | 63                | 62.5              | 62.5             | 62.5              |
| Result<br>No.                 | 1                 | 2                 | ٣                 | 4                | S                 | 9                | 7                | 8                 | 6                 | 10                | 11                | 12              | 13              | 14                | 15                | 16                | 17                | 18               | 19                | 20                | 21                | 22                | 23              | 24                | .25               | 26               | 27                |

| Sequence 18, Appl<br>Sequence 26, Appl<br>Sequence 26, Appl<br>Sequence 68, Appl<br>Sequence 68, Appl<br>Sequence 18, Appl<br>Sequence 18, Appl<br>Sequence 18, Appl<br>Sequence 25, Appl<br>Sequence 25, Appl<br>Sequence 25, Appl<br>Sequence 29, Appl<br>Sequence 29, Appl<br>Sequence 29, Appl<br>Sequence 29, Appl<br>Sequence 147, Appl<br>Sequence 147, Appl<br>Sequence 147, Appl<br>Sequence 147, Appl | Methods for e Containing Fabrics Using Truncated Compositions lenter, P.O. Box 457                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| US-09-081-345-18<br>US-08-470-350B-2<br>US-07-946-421-26<br>US-08-448-905-68<br>US-08-481-985B-68<br>US-08-370-476-68<br>US-08-2017-114-18<br>US-08-505-307-18<br>US-08-505-307-18<br>US-08-505-307-18<br>US-08-51-920-29<br>US-08-361-920-29<br>US-08-361-920-29<br>US-08-488-113B-147<br>US-08-488-113B-147<br>US-08-448-113B-147                                                                             | ALIGNMENTS  ALIGNMENTS  To mam K.  To man Me  To Cellulose  
| 13.3 802<br>13.2 1290<br>13.1 128<br>13.1 307<br>23.1 307<br>23.1 307<br>13.1 316<br>13.1 316<br>13.1 316<br>13.1 316<br>13.0 617<br>13.0 617<br>13.0 617<br>12.9 240<br>12.9 240                                                                                                                                                                                                                               | plication US 366 MATION: T: Farringto Anderson, Glbbs, Mo, Glbbs, Mo, Bergquist Daniels, Morgan, H Williams, INVENTION: TE: PA MINISEE: HOWS EET: Spring TY: Spring HO TYE: PA MINISEE: HOWS EET: Spring HO TYE: PA MINISEE: HOWS TYE: PA MINISEE: TO THE: PA MINISEE: TO THE THE TO THE TO THE TO THE TO THE TO THE TO THE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 28 62.5<br>30 61.5<br>32 61.5<br>32 61.5<br>33 61.5<br>34 61.5<br>36 61.5<br>37 61.5<br>38 61.5<br>44 60.5<br>44 60.5                                                                                                                                                                                                                                                                                           | SULT 1 -09-136-574A Sequence 47, Sequence 47, GENERAL IN TITLE CORRE CORRE TITLE TITLE TITLE TITLE TITLE TITLE TITLE TITLE ORRE TITLE TITLE TITLE TITLE TITLE ATTOR                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|                                                                                                                                                                                                                                                                                                                                                                                                                 | 8 D                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |

QQ ò

```
Containing Fabrics Using Truncated
                                        4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 1426;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Cellulase Enzyme Compositions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Williams, Diane P. TITLE OF INVENTION: Compositions and Methods for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STREET: Spring House Corporate Center, P.O. CITY: Spring House
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 40.7%; Score 191.5; DB 4; Best Local Similarity 43.5%; Pred. No. 9.5e-13; Matches 37; Conservative 15; Mismatches 32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE/DOCKET NUMBER: 1997US001/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,574A
FILING DATE: 19-Aug-1998
CILASSIFICATION: <un control of the co
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Treating Cellulose
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: US 08/932,571
FILING DATE: September 19, 1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY: linear
MOLECULE TYPE: No. 6294366e
SEQUENCE DESCRIPTION: SEQ ID NO: 43:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Howson
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
                                                                                                                                                                                                                                                                                                                                                       Sequence 43, Application US/09136574A
Patent No. 6294866
GENERAL INFORMATION:
APPLICANT: Farrington, Graham K.
                                                                                                                                                                          472 IGASNVTFNFVKLSSGVSGADYYLE 496
                                                                                                                                  64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH: 1426 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Anderson, Paige
Gibbs, Moreland
Bergquist, Peter
Daniels, Roy
Morgan, Hugh W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: Howson and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ZIP: 19477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STATE: PA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF
                                                                                                                                                                                                                                                                                                                      -09-136-574A-43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-136-574A-43
                                                               g
                                                                                                                                  οy
                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Treating Cellulose Containing Fabrics Using Truncated Cellulase Enzyme Compositions
                                                                                                                                                                ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ij
                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                            1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 1751;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Box 457
                                                                                                                                                            34; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Center, P.O.
CITY: Spring House
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Daniels, Roy
Morgan, Hugh W.
Williams, Diane P.
TILE OF INVENTION: Compositions and Methods for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 41.4%; Score 194.5; DB 4;
44.7%; Pred. No. 5.7e-13;
iive 14; Mismatches 32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE/DOCKET NUMBER: 1997US001/CIP
TELECOMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
                                                                                          Score 195.5; DB 4 Pred. No. 1.2e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: US/09/136,574A
                                                                                                                                                         15; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: US 08/932,571 FILING DATE: September 19, 1997 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-09-136-574A-44
                                                                                                                                                                                                                                                                                                                                                           61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                             WAQIGASNVTFNFVKLSSGVSGADYYLE 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REGISTRATION NUMBER: 31,215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: Farrington, Graham K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 44, Application US/09136574A Patent No. 6294366 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE: 19-Aug-1998
CLASSIFICATION: <Unknown>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER: IBM COMPATIBLE OPERATING SYSTEM: DOS CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE CHARACTERISTICS:
LENGTH: 1751 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Anderson, Paige
Gibbs, Moreland
Bergquist, Peter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 44:
                                                                                                41.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NUMBER OF SEQUENCES: 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY: linear MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME: Bak, Mary E
                                                                                                                                                            38; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 44.7⁴
Matches 38; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY: USA
                                                                                          Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ZIP: 19477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 2
US-09-136-574A-44
US-09-136-574A-47
                                                                                                                                                            Matches
```

1;

```
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 595800830 No. 59580083disk of No. 5958083th America, Inc.
STREET: 405 Lexington Avenue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                35.2%; Score 165.5; DB 2; Length 551; 37.9%; Pred. No. 2.2e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ramot University Authority for Applied Research and Industrial Development Ltd. Technion Research and Development Foundation Ltd. Bayer, Edward A. Morag, Ely
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MODIFIED CELLULOSE-BINDING DOMAIN (CBD) PROTEINS AND USE THEREOF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Yeda Research and Development Co. Ltd
                                                                                                                                                                                                                            COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastESG for Windows Version 2.0
CURTENY APPLICATION DATA:
APPLICATION NUMBER: US/09/033,537A
FILING DATE: 02-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3: Browdy and Neimark
419 Seventh Street N.W., Ste. 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE/DOCKET NUMBER: 4492.204-US TELECOMMUNICATION INFORMATION: TELEPHONE: 212-867-0123
                                                                                                                                                                                                                                                                                                                                                            CLASSIFICATION: 008
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 0993/95
FILING DATE: 08-SEP-1995
APPLICATION NUMBER: PCT/DK96/00364
FILING DATE: 03-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
In Stone Washing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | :| |: |: | : | 461 AQIGASNVSAAF--ANFTGSNTDTYVE 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 9, Application PC/TUS9513813
GENERAL INFORMATION:
APPLICANT: Yeda Research and Devel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH: 551 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Wilchek, Meir
Lamed, Raphael
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            33; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Shoham, Yuva]
                                                                                                                                                                  ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEFAX: 212-878-9655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE OF INVENTION:
FITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
                                                                                                                                                 U.S.A.
                                                                                                        New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PCT-US95-13813-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT:
APPLICANT:
                                                                                                                                                 COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3;
                                                                                                                                                                                                                                                                           5916796th America, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 VSGGVKVQYKNND-SAPGDNQIKPGLQLVNTGSSSSVDLSTVTVRYWFTRDGGSSTLVYNC 59
                                                                       Sequence 2, Application US/07862588B
Patent No. 5916796
GENERAL INFORMATION:
APPLICANT: Sch lein, Martin
APPLICANT: Hansen, Christian
APPLICANT: An Enzyme Exhibiting Cellulase Activity
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59167960 No. 5916796disk of No. 5916796th P
STRRET: 405 Lexington Avenue, 62nd floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 700;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          27; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 0S/07/862,588B
FILING DATE: 19920727
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRIOR APPLICATION DATA:
APPLICATION NUMBER: DK 164/90
FILING DATE: 19-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DCT/DK91/00013
FILING DATE: 18-JAN-1991
ATTORNEY ARGENT INFORMATION:
NAME: Zelson, Steve T. / Lambirls, Elias J. REGISTRATION NUMBER: 30,335 / 33,728
REFERENCE/DOCKET NUMBER: 30,335 / 33,728
TELECOMMUNICATION INFORMATION:
TELECHHONE: 212 867 0123
TELEFRAX: 212 867 0129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 1, Application US/09033537A
Patent No. 5958083
GENERAL INFORMATION:
APPLICANT: Onish1, Masahiro
APPLICANT: Fich, Merete
APPLICANT: Toft, Amertte Hanne
APPLICANT: Sh lein, Martin
TITLE OF INVENTION: Prevention Of Back-Staining
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             36.6%; Score 172; DB 2;
44.9%; Pred. No. 5.6e-11;
tive 18; Mismatches 27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ||| :| || :| || :| || :| || 608 DWAKLGGSNIQISFGNHNGA--DSDTYAE 634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       60 DWAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER: IBM PC Compatible
                                                                                                                                                                                                                                                                                                                                                                                                                              Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE CHARACTERISTICS:
LENGTH: 700 amino acids
TYPE: AMINO ACID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 36.6%
Best Local Similarity 44.9%
Matches 40; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELEFAX: 212 867 0298 INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; MOLECULE TYPE: protein US-07-862-588B-2
                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                       STATE: New York COUNTRY: U.S.A. ZIP: 10017
                                                                                                                                                                                                                                                                                                                                                                                                                            MEDIUM TYPE:
                                                          -07-862-588B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-033-537A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
```

ä

Gaps

Indels

```
331 VSGNLKVEFYNSNPSDTTNSINPQFKVTNTGSSAIDLSKLTLRYYYTVDGQKDQTFW-CD 389
                                            1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 493;
                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: BJORNVAG, MAGS E.
APPLICANT: Lange, Niels E.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Moller, Sornen
APPLICANT: Glad, Sanne O. S.
APPLICANT: Kauppinen, Markus S.
APPLICANT: Kauppinen, Markus S.
APPLICANT: Kongsbak, Lars
TITLE OF INVENTION: No. 6187580el Pectate Lyases
FILER REFERENCE: 5378-200-028
FILER REFERENCE: 5378-210-028
FRIOR PRIOR PELICATION NUMBER: US/09/198,955A
PRIOR APPLICATION NUMBER: 1344/97
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-12-02
PRIOR APPLICATION NUMBER: 60/067,240
PRIOR APPLICATION NUMBER: 09/073,684
PRIOR FILING DATE: 1998-05-06
PRIOR RELING DATE: 1998-11-02
PRIOR FILING DATE: 1998-11-02
PRIOR FILING DATE: 1998-11-02
29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               35.1%; Score 165; DB 4; ilarity 36.1%; Pred. No. 2.1e-10; Conservative 23; Mismatches 29
                                                                                                                                                               390 HAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLE 426
                                                                                                                                      61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
23; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQ ID NOS: 32
FastSEQ for Windows Version 4.0
                                                                                                                                                                                                                                                                        US-09-198-955A-12
; Sequence 12, Application US/09198955A
; Patent No. 6187580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM: Clostridium thermocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: Joergensen, Per Linaa APPLICANT: Sch lein, Martin AppLICANT: Hansen, Christian TITLE OF INVENTION: An Enzyme EX) NUMBER OF SEQUENCES: 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 7, Application US/07862588B Patent No. 5916796
                                                                                                                                                                                                                                                                                                                                                                          Andersen, Lene N.
Schulein, Martin
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 35; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NUMBER OF SEQ ID NOS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-198-955A-12
35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-07-862-588B-7
                                                                                        qq
                                                                                                                                         δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Lange, Wiels Erik K.
APPLICANT: Lange, Wiels Erik K.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Bjornvad, Mads E.
TITLE OF INVENTION: Licheniformis
FILE REFRERENES: 5377-200-US
CURRENT APPLICATION NUMBER: US/09/198,956
CURRENT FILING DATE: 1998-11-24
EARLIER APPLICATION NUMBER: 1344/97
EARLIER PILING DATE: 1997-11-24
EARLIER PILING DATE: 1997-11-24
EARLIER FILING DATE: 1997-11-24
EARLIER FILING DATE: 1997-11-24
EARLIER FILING DATE: 1997-11-24
SEQ ID NOS: 26
SOFTWARE: FastSEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    35.1%; Score 165; DB 5; Length 167; 36.1%; Pred. No. 5.7e-11; tive 23; Mismatches 29; Indels 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 165; DB 4; Length 493; Pred. No. 2.1e-10;
                                                                                                                                                   SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13813
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                   NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: BAYER=3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
                                                                                                           COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 10, Application US/09198956
Patent No. 6165769
GENERAL INFORMATION:
APPLICANT: Andersen, Lene N.
APPLICANT: Schulein, Martin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; ORGANISM: Bacillus licheniformis US-09-198-956-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    35.1%;
36.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEX: 248633
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
                                                                                                                                                                                                                                                                             ATTORNEY/AGENT INFORMATION:
                                                               COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 36.1%
Matches 35; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; MOLECULE TYPE: peptide PCT-US95-13813-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        linear
                                                                                                                                                                                                                                                     CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRANDEDNESS:
                                                                                                                                                                                                                               FILING DATE:
                                         20004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-198-956-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Óλ
```

g

; ;

```
66141 PENNIE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 11
US-08-460-462-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       7
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59167960 No. 5916796disk of No. 5916796th America, Inc. STREET: 405 Lexington Avenue, 62nd floor CITY: New York STATE: New York COUNTRY: U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                           SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/862,588B
FILING DATE: 19920727
CLASSIFICATION: 435
PIOOR APPLICATION DATA:
APPLICATION NUMBER: D67-700013
FILING DATE: 19-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/DK91/00013
FILING DATE: 18-JAN-1991
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME: Zelson, Steve T. / Lambiris, Elias J. REGISTRATION NUMBER: 30,335 / 33,728
REFERENCE/DOCKET NUMBER: 3425.204-US
TELECOMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 7:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25.6%; Score 120.5; DB 32.9%; Pred. No. 1.9e-05
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: Shoseyov, Oded
APPLICANT: Shplegl, Itai
APPLICANT: Shplegl, Itai
APPLICANT: Goldstein, Marc A.
APPLICANT: Dol, Roy H.
TILE OF INVENTION: CELLULOSE BINDING DOMAIN
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
                                                                                                                                                                                      COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             491 AQIGRTNVLLAF -- ANFTGSNTDTY 513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 2, Application US/08048164A Patent No. 5496934 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         62 AAMGCGNIRASFGSVNPATPTADTY 86
                                                                                                                                                                       Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   : 531 amino acids
AMINO ACID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 32.99
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-862-588B-7
                                                                                                                                               COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New York
: U.S.A.
                                                                                                                                                                       MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY: U. ZIP: 10036
                                                                                                                        10017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-048-164A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
Sequence 2, Application US/08460462
Patent No. 5670623
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Shpiegl, Itai
APPLICANT: Shpiegl, Itai
APPLICANT: Goldstein, Marc A.
APPLICANT: OF INVENTION: METHODS OF USE OF CELLULOSE BINDING DOMAIN PROTEINS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             34; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Ploppy disk
COMPUTER: BM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NATA:
APPLICATION NUMBER: US/08/460,462
FILING DATE: CONCULTENTLY HEREWITH
SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/048,164A
FILING DATE: 14 APR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24.3%; Score 114; DB 1;
29.9%; Pred. No. 2.3e-05;
tive 22; Mismatches 34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             62 AAMGCGN-----IRASF--GSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         62 AGALLGNSYVDNTSKVTANFVKETASP-TSTYDTYVE 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE: concurrently herewith CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/048,164
FILING DATE: 14-APR-1993
ATTORNEY AGENT INFORMATION:
                                                                                                                  ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7809-0
TELECOMMUNICATION INFORMATION:
TELERONE: (212) 790-9090
TELERAX: (212) 869-864/9741
TELERX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     : (212) 790-9090
(212) 869-8864/9741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE/DOCKET NUMBER: 78
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 29.98
Matches 29; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-048-164A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       U.S.A.
                                                                            FILING DATE: 1. CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY: U. ZIP: 10036
```

```
Sequence 2, Application US/08460458
Patent No. 5738984
Patent No. 5738984
GENERAL INFORMATION:
APPLICANT: Shoseyov, Oded
TITLE OF INVENTION: KITS AND METHODS OF DETECTION USING CELLULOSE BINDING DOMAI NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 TSSMSVEFYNSNKSAQTNSITPIIKITNTSDSDLNLNDVKVRYYYTSDGTQGQTFW-CDH 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
  2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                              3 TSSMSVEFYNSNKSAQTNSITPIIKITNTSDSDLNLNDVKVRYYYTSDGTQGQTFW-CDH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 162;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDIUM IIFE: FIDEN GOMPHUEN
COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
SOFTWARE: Patentin Release #1.0, Version #1.25
SOFTWARE: Patentin Release #1.0, Version #1.25
CURREWT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,458
FILING DATE: CONCURTENTLY herewith
CLASSIFICATION NUMBER: US 08/048,164
FILING DATE: 14-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: MISTOCK, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7809-007
TELEPRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7809-007
TELEPRATION NUMBER: 18,872
REFERENCE/OMMUNICATION INFORMATION:
TELEPRATIC (212) 790-9090
TELERAX: (212) 790-9090
TELERAX: (212) 790-9090
TELERAX: CALAI PENNIE
INFORMATION FOR SEO ID NO: SEOUENCE CHRRACTERISTICS:
LENGTH: 1162 Amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               24.3%; Score 114; DB 1;
29.9%; Pred. No. 2.3e-05;
tive 22; Mismatches 34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                62 AAMGCGN-----IRASF--GSVNPATPTADTYLQ 88
                                                                                                                       62 AGALLGNSYVDNTSKVTANFVKETASP-TSTYDTYVE 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62 AAMGCGN-----IRASF--GSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 2, Application US/08460455
Patent No. 5837814
GENERAL INFORMATION:
APPLICANT: Shoseyov, Oded
APPLICANT: Shoseyov, Itai
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 29.9%
Matches 29, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 , MOLECULE TYPE: protein US-08-460-458-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY: U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY:
                                                                                                                                                                                                                                    US-08-460-458-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-460-455-2
                                         g
                                                                                            δλ
                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                            4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4
                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                   2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          34; Indels 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: Shosyov, Oded
APPLICANT: Shpiegl, Itai
APPLICANT: Shpiegl, Itai
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
TITLE OF INVENTION: CELLULOSE BINDING DOMAIN FUSION PROTEINS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITTL New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         24.3%; Score 114; DB 1; Length 162; 29.9%; Pred. No. 2.3e-05; Live 22; Mismatches 34; Indels
                                                                                                                                                                                                              Length 162;
                                                                                                                                                                                                                                                          34; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
ADDITION TABLECATION DATA:
                                                                                                                                                                                                         Query Match 24.3%; Score 114; DB 1; Best Local Similarity 29.9%; Pred. No. 2.3e-05; Matches 29; Conservative 22; Mismatches 34;
                                                                                                                                                                                                                                                                                                                                                                                                                            62 AGALLGNSYVDNTSKVTANFVKETASP-TSTYDTYVE 97
                                                                                                                                                                                                                                                                                                                                                                                                  62 AAMGCGN-----IRASF--GSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME: MISTOCK, S. Leslie
REGISCRATION NUMBER: 18,872
REFERENCE/CDOCKET NUMBER: 7809-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER: US/08/460,457 FILING DATE: concurrently herewith
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/048,164
FILING DATE: 14-ARR-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; Sequence 2, Application US/08460457; Patent No. 5719044
                     SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                162 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29; Conservative
INFORMATION FOR SEQ ID NO:
                                                                                                          , MOLECULE TYPE: protein US-08-460-462-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ) MOLECULE TYPE: protein US-08-460-457-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
Matches 29; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     New York : U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GENERAL INFORMATION:
APPLICANT: Shoseyo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE: CC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STATE: Ne COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH:
                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

4

```
completed: July
ne: 63 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Search com
Job time:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 24.3%; Score 114; DB 2; Length 162; Best Local Similarity 29.9%; Pred. No. 2.3e-05; Matches 29; Conservative 22; Mismatches 34; Indels
APPLICANT: Goldstein, Marc A.
APPLICANT: Doi, Roy H.
TITLE OF INVENTION: CELLULOSE BINDING DOMAIN PROTEINS
WUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: Shoseyov, oded
APPLICANT: Yosef, Karmey
APPLICANT: Yosef, Karmey
APPLICANT: Shidegi, Itai
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Doi, Roy H.
TITLE OF INVENTION: METHODS OF DETECTION USING THE
TITLE OF INVENTION: CELLULOSE BINDING DOMAIN
NUMBER OF SEQUENCES:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
                                                                                                                                                                                                                                                                                                                                                          CURRENT APPLICATION BELEASE #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460.455
FILING DATE: CONCURTENTLY herewith
CLASSIELCATION NUMBER: US/08/08/164
FILING APPLICATION DATA:
APPLICATION NUMBER: US 08/048/164
FILING APPLICATION NUMBER: US 08/048/164
FILING DATE: H-APR-1993
ATTORNEY/AGENT INFORMATION:
NAMME: Mistorck, S. Leslie
REGISTRATION NUMBER: 7809-009
TELEPHONE: (212) 790-990
TELEPHONE: SEQ ID NO: 2:
CONTRACTOR NO: SEQ ID NO: 2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        62 AGALLGNSYVDNTSKVTANFVKETASP-TSTYDTYVE 97
                                                                                                                           ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
                                                                                                                                                                                                                                           ZIP: 10036
COMPOTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC_DOS/MS_DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 2, Application US/08330394A Patent No. 5856201 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; TOPOLOGY: linear
) MOLECULE TYPE: protein
US-08-460-455-2
                                                                                                                                            STREET: 1155 Aver
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            E: NEW YORK
TRY: U.S.A.
10036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-330-394A-2
```

g ò

```
2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 TSSMSVEFYNSNKSAQTNSITPIIKIINTSDSDLNLNDVKVRYYYTSDGTQCQTFW-CDH 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 162;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            34; Indels
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA: PS/08/330,394A
FILING DATE: 27-CT-1994
FILING DATE: 14-APR-1994
FILING DATE: 14-APR-1994
FILING DATE: 14-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: MISROCK, S. LESLIE
REFERENCE/POCKET NUMBER: 7809-005
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEFAK: (212) 869-8864
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match

24.3%; Score 114; DB 2;
Best Local Similarity 29.9%; Pred. No. 2.3e-05;
Matches 29; Conservative 22; Mismatches 34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2, 2002, 09:12:55
                                                                                                                                                                                                                                                                                                                                                                                               TELEFAX: (212) oc. ...
TELEFAX: (212) oc. ...
TELEX: 66441 PENNIE
TELEX: 66441 PENNIE
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-330-394A-2
```

4 ;

. . . .

```
C;Genetics:
                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                         cellulase (EC 3.2.
cellulase (EC 3.2.
cellulase (EC 3.2.
mannan endo-1,4-be
cellulase (EC 3.2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            scaffolding protei
probably celluloso
scaffolding protei
cellulose 1,4-beta
cellulose-binding
                                                        2, 2002, 09:14:12 ; Search time 54.67 Seconds (without alignments) 156.429 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            endoglucanase - Er
cellulase (EC 3.2.
thermoactive cellu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              hypothetical prote
hypothetical prote
hypothetical prote
                                                                                                                                                                                                                                                                                                                                                                                                                                         1,4-beta-glucanase
xylanase - Caldice
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    protein - Clo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              fami]
                                                                                                                                                                                                                                                                                                                                             Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             probable RTX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      hypothetical cell surface
                                                                                                                                                                                                                                                                                                                                                                                                                     Description
                                                                                                                    1 VSGGVKVQYKNNDSAPGDNQ.....RASFGSVNPATPTADTYLQX
                                                                                                                                                                                                                                                                                                                                                                                                                                                    xylanase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     cipA
          Compugen Ltd
                                                                                                                                                                                         Total number of hits satisfying chosen parameters:
GenCore version 4.5
Copyright (c) 1993 - 2000 Comp
                                                                                                                                                                      283138 seqs, 96089334 residues
                                                                                                                                                                                                                                                                                                                                                                                       SUMMARIES
                                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                        OM protein - protein search, using sw model
                                                                                                                                                   Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                       T31337
T31085
G69593
A26874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JN0111
A48954
T17120
A27198
A43802
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     I40548
B41897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A47704
S36859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             S39962
S54744
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 S12021
PC6006
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A44140
T08685
T25478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 S02711
A41897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          S47466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C97012
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                T30433
                                                                                                                                                                                                             Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                 US-09-917-376-4
                                                                                                                                                                                                                                                                                                                                                                                                                    DB
                                                                                                                                                                                                                                                                                                                                                                                                                    Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          586
1483
1162
1230
                                                                                                                                                                                                                                                                                   PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*
                                                                                                                                          BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                          Query
Match
                                                           July
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        186.5
186.5
186.5
182.5
182.5
180.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               172
165.5
165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             149.5
147.5
141.5
130.5
126.5
                                                                                                             Perfect score:
                                                                                                                                          Scoring table:
                                                                                                                      Sequence:
                                                                                                                                                                      Searched:
                                                                                                                                                                                                                                                                                     Database
                                                          Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Title:
                                                                                                                                                                                                                                                                                                                                                                                                          Result
```

| hypothetical prote | thaumatin-like pro | hypothetical prote | ser-lys rich hypot | polyprotein - hepa | probable purH prot | probable reverse g | hypothetical prote | hensin - rabbit | hemolysin - Aquife | sanko - human | endo-beta-1,3-1,4- | replicative DNA he | protein-tyrosine-p | hypothetical prote | Ig heavy chain V r |  |
|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|-----------------|--------------------|---------------|--------------------|--------------------|--------------------|--------------------|--------------------|--|
| H72684             | JC7201             | T47790             | T40462             | T01075             | C70717             | G72614             | 850669             | T30549          | H70393             | A59386        | н95976             | AB2512             | B44390             | AH2515             | E37267             |  |
| ~                  | 7                  | N                  | ~                  | ď                  | ~                  | N                  | ~                  | ~               | ~                  | ~             | 7                  | N                  | 7                  | ~                  | 7                  |  |
| 271                | 247                | 535                | 751                | 1435               | 523                | 1222               | 1571               | 1594            | 398                | 2403          | 269.               | 454                | 802                | 4936               | 122                |  |
| 14.3               | 14.1               | 14.0               | 13.9               | 13.9               | 13.8               | 13.6               | 13.6               | 13.6            | 13.5               | 13.5          | 13.3               | 13.3               | 13.3               | 13.3               | 13.2               |  |
| 67                 | 66.5               | 99                 | 65.5               | 65.5               | 65                 | 64                 | 64                 | 64              | 63.5               | 63.5          | 62.5               | 62.5               | 62.5               | 62.5               | 62                 |  |
| 30                 | 31                 | 32                 | 33                 | 34                 | 35                 | 36                 | 37                 | 38              | 39                 | 40            | 41                 | 42                 | 43                 | 44                 | 45                 |  |
|                    |                    |                    |                    |                    |                    |                    |                    |                 |                    |               |                    |                    |                    |                    |                    |  |

## ALIGNMENTS

```
TASSULE 1
TASSULE 1
TASSULE 1
TASSULE 1
C. Species: Anaerocellum thermophilum (fragment)
C. Species: Anaerocellum thermophilum (fragment)
C. Species: Anaerocellum thermophilum
C. Accession: T31337
TACCESSION: T31337
TACCESSION: T31337
TATILE: Properties and gene structure of a bifunctional cellulolytic enzyme (CelA) f
M. Accession: T31337
TATILE: Properties and gene structure of a bifunctional cellulolytic enzyme (CelA) f
M. Accession: T3137
TACCESSION: T3137
TACCESSION: T3137
TACCESSION: T3137
TACCESSION: MUD: 98154434
TACCESSION: T3137
TACCESION: T3137
```

of the signal region of Bacillus subitil

```
4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA 62
    A; Accession: S49103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Callulase (EC 3.2.14) bglC precursor - Bacillus subtilis

NiAlternate names: endo-1.4-beta-glucanase
C;Species: Bacillus subtilis
N;Alternate names: endo-1.4-beta-glucanase
C;Species: Bacillus subtilis
C;Date: Ob-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
C;Accession: G65953; Az6114; I40353; S24239; S49103; I39803
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G; Azevedo, V.; Berter
C; Bron, S; Brouillat, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chc
A; Ebrilch, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Febret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Laubber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Lapidus, S.; Maucell
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Perro, V.; Pohl, T.M.; Pothotes, S.; Schleich, S.; Schroeter, R.; Socfone, F.; Sekiguchi, J.; Sekowska, A.; Scrone,
A;Authors: Schleich, S.; Schroeter, R.; Socfone, F.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipet, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, R.; Attle: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Aterence number: A69580; MUID:98044033
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-508 kRUN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Molecule type: DNA
A; Residues: 10-291, 'N', 293-508 <LIN2>
A; Cross-references: EMBL: K67044; NID: g39776; PIDN: CAA47429.1; PID: g39777
A; Experimental source: strain CK-2
R; Wolf, M.; Geczi, A.; Borriss, R.
submitted to the EMBL Data Library, December 1993
A; Description: Genes encoding beta glucan-hydrolyzing enzymes in Bacillus subtilis; cons
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A.Cross-references: GB: 299113; GB: AL009126; NID: 92634090; PIDN: CAB13696.1; PID: e1183471; A.Cross-references: Strain 168
R. Experimental source: strain 168
R. MacKay, R.M.; Lo, A.; Willick, G.; Zuker, M.; Baird, S.; Dove, M.; Moranelli, F.; Seli
Nucleic Acids Res. 14, 9159-9170, 1986
A.Title: Structure of a Bacillus subtilis endo-beta-1,4-glucanase gene.
A.Reference number: A26114; MUID: 87066783
A.Accession: A26114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        part of this sequence, including the amino end of the mature form, was confirmed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Note: part of this sequence, including the amino end of the mature form, was confirme A; Note: part of this sequence, including the amino end of the mature form, was confirme. An thindahl, V.; Aa, K.; Tronsmo, A.
Antonie Van Leeuwenhoek 66, 327-32, 1994
A; Title: Nucleotide sequence of an endo-beta-1, 4-glucanase gene from Bacillus subtilis and Reference number: 140353; MUID:95225656
A; Accession: 140353
A; Residues: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Molecule type: DNA
A; Residues: 10-291, VV, 293-508 < LINI>
A; Residues: 10-291, VV, 293-508 < LINI>
A; Cross-references: EMBL:X67044; NID:939776; PIDN:CAA47429.1; PID:939777
A; Cross-reference number: S24239
A; Recession: S24239
                                                                                  1;
                                                                                                                                                                                                      Gaps
                                                                                                                                                            63
                                                                                                                                                        4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: DNA
A; Residues: 10-508 <MAC>
A; Cross-references: GB:X04689; NID:g39823; PIDN:CAA28392.1; PID:g39824
A; Experimental source: strain PAP115
    Length 1779;
    DB 2;
                                         .4e-11;
                                                                              14; Mismatches
Score 190.5;
Pred. No. 1.4
                                                                                                                                                                                                                                                                                                                                                                                               1172 IGASNVTFNFVKLTSGVSGADYYLE 1196
                                                                                                                                                                                                                                                                                                                 64 MGCGNIRASFGSVNPATPTADTYLQ 88
40.5%;
                                                                              Conservative
                                         Similarity
    Query Match
                                                 Local
                                                                              Matches
                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                           윱
                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                 ò
```

```
A;Pathway: cellulose degradation
C;Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
F;1-38/Domain: signal sequence #status predicted <SIG>
F;39-508/Product: cellulase #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Pathway: cellulose degradation
C;Reywords: alternative initiators; extracellular protein; glycosidase; hydrolase; po
F;l-38/Domain: (or 2-38) signal sequence #status predicted <SIG>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A Molecule type: DNA
A Residues: 1-508 < ROB1>
A Cross-references: GB:M16185, NID:g143007; PIDN:AAA22496.1; PID:g143008
A) Cross-references: Strain DLG
A) Accession: B26874
A; Molecule type: protein
A) Residues: 39-53 < ROB2>
A) Experimental source: strain DLG
A) Residues: 39-53 < ROB2>
A) Experimental source: strain DLG
A) Comment: The low molecular weight of the mature protein suggests carboxyl-terminal
C) Function:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A)Status: preliminary; translated from GB/EMBL/DDBJ
A)Molecule type: DNA
A)Rolecule type: DNA
A)Residues: 10-13, VV, 17-18, VV, 20-21, F', 23, A, 25-26, AI', 29-31, F', 33, PQ'
A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Residues: 10-13, VV, A)Resi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Cellulase (EC 3.2.1.4) precursor - Bacillus subtilis (strain DLG)
N;Alternate names: endo-1,4-beta-glucanase
C;Species: Bacillus subtilis
C;Species: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 15-Oct-1999
A; Wolecule type: DNA
A; Residues: 10-508 < WOL>
A; Cross -refrences: EMBL: 229076; NID:9509266; PIDN:CAA82317.1; PID:9509267
A; Seo, Y.S.; Lee, Y.H.; Pek, U.H.; Kang, H.
Korean J. Microbiol. 24, 236-242, 1986
A; Title: Analysis on the nucleotide sequence of the signal region of Bacilla, A; Reference number: 139803
A; Reference number: 139803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 508;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C;Accession: A26874; B26874
R;Robson, L.M.; Chambliss, G.H.
A; Tittle: Endo-betra-1-4-q1ucanase gene of Bacillus subtilis
A;Reference number: A26874; MUID:87194581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 40.4%; Score 190; DB 2; L
Best Local Similarity 40.7%; Pred. No. 4.3e-12;
Matches 35; Conservative 21; Mismatches 26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            40.4%; Score 190; DB 2; 43.0%; Pred. No. 4.3e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match

Query Match

40.4%; Score 190; DB

Best Local Similarity 43.0%; Pred. No. 4.3e-
Matches 37; Conservative 17; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   :||||: | ||||: 419 OIGCGNVTHKFVTLHKPKQGADTYLE 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  63 AMCCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Accession: A26874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Gene: bglC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C; Function:
```

'n

Gaps

Indels

C; Keywords: glycosidase; hydrolase; polysaccharide degradation

3

1;

g

셤 δ

```
cellulase (EC 3.2.1.*) precursor, thermoactive - Caldocellum saccharolyticum C; Species: Caldocellum saccharolyticum C; Species: Caldocellum saccharolyticum C; Species: Caldocellum saccharolyticum C; Species: Tariology #sequence_revision 15-Oct-1999 #text_change 17-Nov-2000 C; Accession: T17120; A43745

R; TeVO, V. S.; Saul, D.J.; Bergquist, P.L. Appl. Microbiol. Biotechnol. 43, 291-296, 1995
A; Title: CalA, another gene coding for a multidomain cellulase from the extreme therm A; Reference number: 218698; MUID:95336703
A; Accession: T17120
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Status: DNA
A; Residues: 1-1742 < TEO>
A; Cross-references: EMBL:132742; NID:9537499; PID:9537500; PIDN:AAA91086.1
R; Luethl, E.; Bhana Jasmat, N; Grayling, R.A.; Love, D.R.; Bergquist, P.L.
A; Title: Cloning, sequence analysis, and expression in Escherichia coli of a gene cod A; Reference number: A43745; MUID:91247819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Wildrase (EC 3.2.1.4) precursor - Bacillus subtilis (strain IF03034)
N;Alternate names: endo-1,4-beta-glucanase
C;Species: Bacillus subtilis
C;Species: Bacillus subtilis
C;Accession: A27198
R;Nakamura, A.; Uozumi, T.; Beppu, T.
Eur. J. Blochem. 164, 317-320, 1987
A;Title: Nucleotide sequence of a cellulase gene of Bacillus subtilis.
A;Reference number: A27198; MUID:87190397
                                                                                                                                                                                                                                                    A;Residues: 1516-1544,'A',1546-1742 <LUE>
A;Cross-references: EMBL:M36063; NID:g144292; PIDN:AAA72860.1; PID:g144293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                    SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Molecule type: DNA
A;Residues: 1-499 <NAK>
A;Cross-references: GB:M28332; NID:g142670; PIDN:AAA22307.1; PID:g142671
                                                                                                                                                             1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 1742;
                                                                                                Length 1331;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                          Indels
                                                                      Score 186.5; DB 2;
Pred. No. 2.8e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; Score 186.5; DB 2; Pred. No. 3.7e-11; 18; Mismatches 31
                                                                                                                                                          18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                  62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C; Keywords: glycosidase; hydrolase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           39.78;
                                                                                                39.7%;
                                                                                   Query Match
Best Local Similarity 42.5%
Matches 37, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 42.5%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: DNA
A; Residues: 1516-1544,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Accession: A43745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Gene: celA
                                                                                                                                                                                                                       7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          C;Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C)
                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                        C. Accession: JN0111

R.Park, S.H.; Kim, H.K.; Pack, M.Y.

Agiic. Blol. Chem. 55, 441-448, 1991

A.Title: Characterization and structure of the cellulase gene of Bacillus subtilis BSE61

A.Reference number: JN0111; MUID:91299280

A.Accession: JN0111

A.Accession: JN0111

A.Accession: JN0111

A.Accession: JN0111

A.Residues: 1-499 < PARN

A.Residues: 1-499 < PARN

A.Residues: 1-499 < PARN

A.Cross-references: GB:D01057; NID:9216387; PIDN:BAA00859.1; PID:d1001323; PID:9216388

A.Note: the authors translated the codon ATA for residue 102 as Tyr

C.Comment: The low molecular weight of the mature protein suggests carboxyl-terminal pro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Molecule type: nucleic acid
A; Residues: 1-1331 (GIBA)
A; Residues: 1-1331 (GIBA)
A; Residues: 1-1331 (GIBA)
A; Cross-references: GB:L01257; NID:g144290; PIDN:AAA71887.1; PID:g144291
A; Note: sequence extracted from NCBI backbone (NCBIR:121576, NCBIP:121577)
B; Luethi, E.; Bhana Jasmat, N.; Grayling, R.A.; Love, D.R.; Bergquist, P.L.
A; Little: Cloning, sequence analysis, and expression in Escherichia coli of a gene coding
A; Reference number: A43745; MUID:91247819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C;Accession: A48954; B3745
R;Gibbs, M.D.; Saul, D.J.; Luthl, E.; Bergquist, P.L.
Appl. Environ. Microbiol. 58, 3864-3867, 1992
A;Title: The Determannanase from "Caldocellum saccharolyticum" is part of a multidomain A;Reference number: A48954; MUID:93119139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Pathway: cellulose degradation
C;Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
F;1-29/Domain: signal sequence #status predicted <SIG>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  mannan endo-1,4-beta-mannosidase (EC 3.2.1.78) - Caldocellum saccharolyticum
                                                                                                                                                                                                                                                                                                                                                                     Species: Bacillus subtilis
.Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 28-May-1999
A;Residues: 1-337, 'PPROHOHRQ' <LUE>
A;Cross-references: EMBL:M36063; NID:g144292; PIDN:AAA72861.1; PID:g144294
A;Note: the authors translated the codon CAC for residue 262 as Glu
A;Note: this sequence has been revised in reference A48954
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                        cellulase (EC 3.2.1.4) precursor - Bacillus subtilis (strain BSE616)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 499,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 40.0%; Score 188; DB 2; Length 49 Best Local Similarity 40.0%; Pred. No. 6.7e-12; Matches 34; Conservative 18; Mismatches 31; Indels
                                                                                                                                                                                                                                                                                                                                        N;Alternate names: endo-1,4-beta-glucanase
C;Species: Bacillus subtilis
C;Date: 17-Apr-1993 #sequence_revision 17-P
                                                                                                                                419 QMGCGNLTHKFVTLHKPKQGADTYLE 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           :||||: | ::: ||||: |
LGCGNVTYFFVTLHKPKQGADTYLE 435
                                                                                      63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N;Alternate names: beta-mannanase
C;Species: Caldocellum saccharolyticum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           411
```

ò qq

õ 쉽 ä

ij

Gaps

;;

Length 1039;

DB 2;

```
hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C;Species: Bacillus lautus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 15-Oct-1999
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 15-Oct-1999
C;Accession: A41897; S27498
R;Hansen, C. K.; Diderichsen, B.; Jorgensen, P.L.
J. Bacteriol. 174, 3522-3531, 1992
A;Title: celA from Bacillus lautus PL236 encodes a novel cellulose-binding endo-beta-A;Reference number: A41897; MUID:92276330
A;Reference number: A41897
A;Refer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Diffunctional cellulase precursor - Bacillus sp.
C;Species: Bacillus sp.
C;Accession: I40548
R;Han, S;J; Yoo, Y;J; Kang, H;S.
J; Biol, Chem. 270, 26012-26019, 1995
A;Title: Characterization of a bifunctional cellulase and its structural gen A;Reference number: I40548; MuID:96029707
A;Accession: I40548
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Residues: 1-486 <RES>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Cross-references: EMBL:M76588; NID:g142661; PIDN:AAA22302.1; PID:g142662 A;Experimental source: PL236
A;Note: sequence extracted from NCBI backbone (NCBIP:104604)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Cross-references: EMBL: U27084; NID: 9857575; PIDN: AAC43478.1; PID: 9857576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7 VQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMGC 66
A;Cross-references: EMBL:X13602; NID:g40645; PIDN:CAA31936.1; PID:g40646 (C;Genetics: celB A;Gene: celB A;Gene: celB C;Function: A;Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucosition: hydrolysis of 1,4-beta-Tylanase A homology Superfamily: Streptomyces endo-1,4-beta-xylanase A homology C;Keywords: glycosidase; hydrolase; polysaccharide degradation F;1-28/Domain: signal sequence #status predicted csic>F;29-1039/Product: cellulase #status predicted csic>F;72-373/Domain: Streptomyces endo-1,4-beta-xylanase A homology csxx>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              38.4%; Score 180.5; DB 2 43.9%; Pred. No. 1.1e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 38.8%; Score 182.5; DB 2
Best Local Similarity 41.9%; Pred. No. 5.6e-11;
Matches 36; Conservative 18; Mismatches 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           cellulase homolog - Bacillus lautus (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       479 QIGASNVTFKFVKLSSSVSGADYYLE 504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SNIQISFG--NHTGTNSDTYVE 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          67 GNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  36; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                            S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        #43802

Cellulase (EC 3.2.1.4) / cellulose 1.4-beta-cellobiosidase (EC 3.2.1.91) - Caldocellum s
Cellulase (EC 3.2.1.4) / cellulose 1.4-beta-cellobiosidase (EC 3.2.1.91) - Caldocellum s
Cellulase (EC 3.2.1.91) - Caldocellum saccharolyticum
C; Species: Caldocellum saccharolyticum
C; Date: 30-Jan-1993 #sequence_revision 30-Sep-1993 #text_change 10-Jul-1998
C; Accession: A43802
C; Accession: A43802
C; Accession: A43802
C; Accession: A3802
C; Acc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               tra
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            exc
A; Experimental source: strain IFO3034
C; Function:
A; Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as A; Pestivay: cellulose degradation
C; Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
C; Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
F; 1-36/Domain: signal sequence #status predicted <SIG>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Accession: A43802
A; Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Title: Nucleotide sequence of a gene from Caldocellum saccharolyticum encoding for A;Reference number: S02711; MUID:89098398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N'Alternate names: endo-1.4-beta-glucanase
N'Contains: cellulase (EC 3.2.1.4); cellulose 1,4-beta-cellobiosidase (EC 3.2.1.91)
C;Species: Caldocellum saccharolyticum
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Oct-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   . ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Pathway: celulose degradation
C; Superfamily: Streptomyces endo-1,4-beta-xylanase A homology
C; Keywords: glycosidase; hydrolase; polysaccharide degradation
F; 20-320/Domain: Streptomyces endo-1,4-beta-xylanase A homology <SXX>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ÷
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              L.W.; Bergquist,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 2; Length 915;
                                                                                                                                                                                                                                                                                                                                                                                                               Length 499;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 cellulase (EC 3.2.1.4) precursor - Caldocellum saccharolyticum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                               39.6%; Score 186; DB 2;
illarity 40.7%; Pred. No. 1.1e-11;
Conservative 19; Mismatches 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     38.8%; Score 182.5; DB 2; 41.9%; Pred. No. 4.8e-11; iive 18; Mismatches 31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     C; Accession: $02711
R; Saul, D.J.; Williams, L.C.; Love, D.R.; Chamley,
Nucleic Acids Res. 17, 439, 1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Reference number: A43802; MUID: 91136262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   410 QIGCGNLTHKFVTLHKPKQGADTYLE 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QIGASNVTFKFVKLSSSVSGADYYLE 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A, Residues: 1-915 <SAU>
A, Cross-references: EMBL: X13602
C, Function:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity 41.9%
les 36; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Accession: S02711
A; Molecule type: DNA
A; Residues: 1-1039 <SAU>
                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 35; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Si
Matches 36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          å
```

ç

2

Gaps

.. 3

Length 145; Indels

DB 2;

structural gene: the

ä

ò q ò QQ

```
Signature clostridium thermocellum
Ni Alternate names: probable cellulosome protein large chain SL
CipAcession: Signetistium thermocellum
C; Species: Clostridium thermocellum
C; Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change Ol-Dec-2000
C; Accession: Signetistium thermocellum
C; Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change Ol-Dec-2000
C; Accession: Signetistium thermocellum
C; Accession: Signetistium A.L.
Submitted to the EMBL Data Library, January 1993
A; Reference number: Signetistium A.L.
A; Residues: preliminary
A; Molecule type: DNA
A; Residues: 1-1854 <GER>
A; Cross-references: EMBL:L08665
A; Cross-references: EMBL:L08665
A; Cross-reference: Signetistium thermocellum gene (cipA) encoding the cellulosom
A; Reference number: Signetistium thermocellum gene (cipA) encoding the cellulosom
A; Reference number: Signetistium thermocellum gene (cipA)
A; Residues: 1-1615,1617-1854 <GEZ>
A; Cross-references: EMBL:L08665
A; Title: Cloning of a Clostridium thermocellum DNA fragment encoding polypeptides tha
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Molecule type: DNA
A; Residues: 1216-1232, 'Y', 1234-1241, 'T', 1243-1319, 'R', 1321-1615, 1617-1768, 'R', 1770-18
A; Cross-references: EMBL:X67406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             R.Fujino, T.; Beguin, P.; Aubert, J.P.
J. Bacteriol. 175, 1891-1899, 1993
A,Title: Organization of a Clostridium thermocellum gene cluster encoding the cellulo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Molecule type: DNA
A;Residues: 1216-1232,'Y',1234-1241,'T',1243-1319,'R',1321-1615,1617-1854 <FUJ>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Cross-references: EMBL: X67506; NID: 9296879; PID: 9296880; PIDN: CAA47840.1
C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         365 VSGNLKVEFYNSNPSDJTNSINPOFKVTNTGSSAIDLSKLTLRYYYTVDGOKDOTFW-CD 423
                                                                                                                                                                                                          36 IKGEVVLQYANGNAGATSNSINPRFKIINNGTKAINLSDVKIRYYYTKEGGASQNFW-CD 794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10;
Length 879;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 1854;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               35.1%; Score 165; DB 2; Length 185
36.1%; Pred. No. 6.9e-09;
iive 23; Mismatches 29; Indels
                                                                                    Indels
        DB 2;
35.2%; Score 165.5; DB 2; 33.0%; Pred. No. 2.7e-09; ive 26; Mismatches 32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RiBeguin, P. Submitted to the EMBL Data Library, August 1992 A; Reference number: S28659 A; Accession: S28659
                                                                                                                                                                                                                                                                                                                                                                                                        795 WSSAGNSNVTGNFFNLSSPKEGADTCLE 822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Reference number: 218847; MUID: 93209931
                                                                                                                                                                                                                                                                                                                                61 WAAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Cross-references: EMBL: X67406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 36.1%
Matches. 35; Conservative
                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Molecule type: DNA
A; Residues: 1821-1854 <FU2>
                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Accession: S25767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Accession: T18261
        Query Match
                                                      Local
                                              Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       pp
                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δλ
                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      National and an entires; entired and an entir
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C; Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 15-Oct-1999
C; Accession: A47704
F; Mazlewood, G-P., Davidson, K.; Laurie, J.I.; Huskisson, N.S.; Gilbert, H.J.
J. Gen. Microbiol. 139, 307-316, 1993
A; Title: Gene sequence and properties of Cell, a family E endoglucanase from Clostridium
A; Reference number: A47704; MUID:93171873
                                                                                                                                                                                                                                                ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Status: preliminary
A;Molecule type: DNA; protein
A;Residues: 1-879 <HAZ.
A;Cross-references: GB:L04735; NID:g144807; PIDN:AAA20892.1; PID:g144808
A;Note: sequence extracted from NCBI backbone (NCBIN:125637, NCBIP:125638)
C;Keywords: glycosidase; hydrolase
                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 VSGGVKVQYKNND-SAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNC 59
                                                                                                                                                                                                                                                                                                                            4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4;
                                                                                                                                                                                                                                                5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 700;
                                                                                                                                                                    Length 486;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27; Indels
                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          endoglucanase I (EC 3.2.1.-) CelI - Clostridium thermocellum C; Species: Clostridium thermocellum
                                                                                                                                                                                                                                            32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   36.6%; Score 172; DB 2; ilarity 44.9%; Pred. No. 4.5e-10; Conservative 18; Mismatches 27,
                                                                                                                                                            Query Match 38.3%; Score 180; DB 2; Best Local Similarity 38.8%; Pred. No. 4.5e-11; Matches 33; Conservative 18; Mismatches 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          608 DWAKLGGSNIQISFGNHNGA--DSDTYAE 634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            60 DWAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N; Alternate names: endo-1, 4-beta-glucanase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          cellulase (EC 3.2.1.4) - Bacillus lautus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    411 LGCGNVSHTVVTLHKPKQGADTYLE 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
Matches 40; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Accession: A47704
    C;Genetics:
A;Gene: cel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               13
```

ò g ò 셤 .; 2

QQ οy

Search completed: July 2, 2002, 09:14:14 Job time: 142 sec

aleuria aur caenorhabdi bacillus su bordetella

microbispor mus musculu aspergillus

escherichia bacillus th

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                              linkages in cellulose.
                                                                                                                                                                                                                                                                                                                                                STRAIN-DLG;
MEDLINE=81194581; PubMed=3106328;
Rebson L.M., Chambliss G.H.;
"Endo-beta-1, 4-glucanase gene of Bacillus subtilis DLG.";
J. Bacteriol. 169:2017-2025(1987).
-: CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PROTON DONOR (BY SIMILARITY).

CELLULOSE-BINDING (BY SIMILARITY).

339DQ4EE95A63EE1 CRC64;
                                                                                                                                                                                                                       01-AUG-1988 (Rel. 08, Created)
01-NOV-1991 (Rel. 20, Last sequence update)
01-NOV-1994 (Rel. 30, Last sentetion update)
Endoglucanase precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
                                                 P34446
P54507
P17740
O75698
P26414
P04945
P06329
P46849
P58127
Q45733
P18891
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; M16185; AAA...

EMBL; M16185; AAA...

PIR; A26874; A26874.

R HSSP; O65465; IA3H.

DR InterPro; IPR001366; CBD_3.

DR Pfam; PF00042; CBD_3; 1.

DR Pfam; PF001947; CBD_3; 1.

DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.

DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.

M Cellulose degradation; Hydrolase; Glycosidase; Signal.

FT SIGNAL

T 29

CHAIN

T 29

CHAIN

T 257

CELULOSPHILE (BY SIMILARITY OF T)

T 257

T 339D04EE95A63EEI CRC64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       40.4%; Score 190; DB 1; Length 499; 43.0%; Pred. No. 4.4e-14;
                                                                                                                                                                                                                                                                                               Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
                                                                                                                                                                                                   499 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             17; Mismatches
                                                                                                                                                ALIGNMENTS
HV50_MOUSE
RTCA_ECOLI
RTCA_ECOS7
C9CA_BACTO
LECF_ALEAU
PAT2_CAEEL
COTN_BACSU
RECA_BORPE
HUG1_HUMAN
                                                                                           GUNA_MICBI
KV6K_MOUSE
XYN1_ASPAW
                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             37; Conservative
                                                                                                                                                                                                    STANDARD;
120
338
342
1157
1226
261
352
352
362
456
108
                                                                                                                                                                                                                      01-AUG-1988 (Rel. 08,
01-NOV-1991 (Rel. 20,
01-OCT-1994 (Rel. 30,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity
                                                                                                                                                                                                                                                                                      Bacillus subtilis.
                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
 13.0
112.9
112.7
112.7
112.7
112.7
112.7
112.7
                                                                                                                                                                                                                                                                                                                                                                                                                                    HYDROLASES).
                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=1423;
                                                                                                                                                                                                                                                                            BGLC OR GLD
                                                                                                                                                                                                                                                                  (Cellulase)
                                                                                                                                                                                                    GUN1_BACSU
 Query Match
                                                                                                                                                                                                              P07983;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
  2, 2002, 09:18:03; Search time 28.05 Seconds (without alignments) 122.853 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           clostridium
clostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            clostridium
clostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     erwinia car
erwinia car
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                clostridium
mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                m bifunctio
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                rhizobium m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    mus musculu
mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                         caldocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      caldocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       paenibacill
paenibacill
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  clostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             erwinia car
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     methanococc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          saccharomyc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      thermoplasm
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               schistosoma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      peroxisom
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               mus musculu
escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              c endogluca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               rhodobacter
                                                                                                                                                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                        Description
                                                                                                                          470
1 VSGGVKVQYKNNDSAPGDNQ.....RASFGSVNPATPTADTYLQX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           002934
001866
001866
0059395
0059395
0059395
0059395
0059316
0017446
0017446
0017446
0017446
0017446
0017446
0017446
0017446
0017446
0017446
                                                                                                                                                                                                                                                                                                                                                                                                                                    P10475 P23549 P22533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P22534
P10474
P29718
P29719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            P03991
P32138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P38108
           4.5
Compugen Ltd
                                                                                                                                                                                                            Potal number of hits satisfying chosen parameters:
                                                                                                                                                                                         105224 seqs, 38719550 residues
                                                                                                                                                                                                                                                                                                                                                                        SUMMARIES
            GenCore version
Copyright (c) 1993 - 2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FLA1_METVA
PUR9_MYCTU
ATC5_YEAST
FLA1_THEVO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HA1Q_MOUSE
HA1W_MOUSE
YIHQ_ECOLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GUX2_CLOSR
GUNZ_CLOSR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CBPA_CLOCL
HV39_MOUSE
                                                                                                                                                                                                                                                                Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GUNA_CALSA
GUNB_CALSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GUNI_CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PSEFL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GUNW_ERWCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TY3H_SCHMA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RHIME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PTN8_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MUCB_PSEAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CEA_PAELA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GUNA_PAELA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CIPA_CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GUNN_ERWCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ERWCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               нтs6_Rноsн
                                                     protein search, using sw model
                                                                                                                                                          BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ICEN_E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          YA42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EXOK
                                                                                                                                                                                                                                 seq length: 0
seq length: 200000000
                                                                                                                US-09-917-376-4
                                                                                                                                                                                                                                                                                                                                                                                          Query
Match Length DB
                                                                                                                                                                                                                                                                                                          SwissProt_40:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                        186.5
186.5
182.5
180.5
165.5
165.5
165.7
165.7
167.5
140.5
141.5
141.5
141.5
141.5
141.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             62.5
62.5
62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61.5
61.5
61.5
61.5
                                                                                                                           Perfect score:
                                                                                                                                                           Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    63.1
                                                     •
                                                                                                                                                                                                                                 Minimum DB
Maximum DB
                                                   OM protein
                                                                                                                                       Sequence:
                                                                                                                                                                                         Searched:
                                                                                                                                                                                                                                                                                                          Database
                                                                        Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Result
                                                                                                                                                                                                                                                                                                                                                                                                      Š
```

Gaps

4

```
ò
                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                               δŏ
                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMIL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to licenseelab-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                           Lindahl V., Aa K., Tronsmo A.;
"Nucleotide sequence of an endo-beta-1,4-glucanase gene from Bacillus
subtilis CK-2.";
             Antonie Van Leeuwenhoek 66:319-326(1994).
-!- CATALTIC ACTIVITY: Endchydrolysis of 1,4-beta-D-glucosidic
linkages in cellulos to CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
-!- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-9225655; PubMed-7710279; MEDLINE-92225655; PubMed-7710279; MEDLINE-92225655; PubMed-7710279; MEDLINE-92225655; PubMed-7710279; Archaracterization of production and enzyme properties of an endo-beta-1,4-glucanase from Bacillus subtilis CK-2 isolated from compost soil.";
4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA
                                                                                                                                                                                                                                                                                                                           MEDLINE-87066783; PubMed-3024130;
Mackay R.M., Lo A., Willick G., Zuker M., Baird S., Dove M.,
Moranelli F., Seligy V.;
"Structure of a Bacillus subtilis endo-beta-1,4-glucanase gene.";
Nucleic Acids Res. 14:9159-9170(1986).
                                                                                                                                                              01-JUL-1989 (Rel. 11, Created)
01-JUL-1989 (Rel. 11, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Endoglucanase precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
(Garboxymethyl-cellulase) (CMCASE) (Cellulase).
BGLC OR GLD OR EGLS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Rose M., Entian K.;
Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                  Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
                                                                                                                                      499 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Antonie Van Leeuwenhoek 66:327-332(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL, X67044; CAA97429.1; -...
EMBL, 27324; CAA97610.1; ALT_INIT.
EMBL, 299113; CAB13696.1; ALT_INIT.
PIR; A26114; A26114.
HSSP; O85465; 1A3H.
                                                                OMGCGNLTHKFVTLHKPKOGADTYLE 435
                                                                                                                                        PRT;
                                                 63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN=CK-2;
MEDLINE=95225656; PubMed=7710280;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; Z29076; CAA82317.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CAA28392.1;
                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 30-45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                          Bacillus subtilis
                                                                                                                                                                                                                                                                             NCBI_TaxID=1423;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HYDROLASES)
                                                                                                                                                                                                                                                                                                                  STRAIN-PAP115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; X04689;
                                                                                                                                        GUNZ_BACSU
P10475;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN=168
                                                                                                               RESULT 2
GUNZ_BACSU
                                                                        410
                       q
                                                                                                                                        ò
```

```
ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Agric. Biol. Chem. 55:441-448(1991).
--- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
--- Linkages in cellulose.
---- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Park S.H., Kim H.K., Pack M.Y.; "Characterization and structure of the cellulase gene of Bacillus subtilis BSE616.";
                                                                                                                                                                                                                                                                        PROTON DONOR (BY SIMILARITY).
NUCLEOPHILE (BY SIMILARITY).
CELLULOSE-BINDING (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-NOV-1991 (Rel. 20, Last sequence update)
01-FEB-1995 (Rel. 31, Last amnotation update)
Endoglucanase precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
(Carboxymethyl-cellulase) (CMCASE) (Cellulase).
                                                                                                                                                                                                                                                                                                                                                                                                                                                       40.4%; Score 190; DB 1; Length 499; 40.7%; Pred. No. 4.4e-14; Live 21; Mismatches 26; Indels
                                                                                                                    Prodom; PD001947; CBD_3; 1.
PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
Cellulose degradation; Hydrolase; Glycosidase; Signal;
                                                                                                                                                                                                                                                                                                                                                     -> N (IN REF. 2).
8F735FF711B3EAE2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            499 AA.
                                                                                                                                                                                                                                                      ENDOGLUCANASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; D01057; BAA00859.1; -.
PIR; J00111; J00111.
HSSP; 086465; 1A3H.
InterPro; IPR001956; CBD_3.
InterPro; IPR001547; Glyco_hydro_F5.
Pfam; PF00942; CBD_3; 1.
Pfam; PF00150; cellulase; 1.
                  InterPro; IPR001956; ČBD_3.
InterPro; IPR001547; Glyco_hydro_F5.
Pfam: PF00942; CBD_3; 1.
Pfam: PF00150; cellulase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               410 QIGCGNVTHKFVTLHKPKQGADTYLE 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-91299280; PubMed-1368694;
                                                                                                                                                                                                                        29
499 EP
169 PI
257 NG
499 CI
283
55287 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1991 (Rel. 20, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
SubtiList; BG10437; bglC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
Labes 35; Conserve
                                                                                                                                                                                                                                                                           169
257
350
283
499 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                      Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bacillus subtilis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=1423;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HYDROLASES)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-BSE616;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GUN3_BACSU
P23549;
                                                                                                                                                                                                                                                                                ACT_SITE
                                                                                                                                                                                                                                                                                                          SILE
                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                              SIGNAL
                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GUN3_BACSU
```

```
ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                             SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GUNA_CALSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ID DT ACC DD TE DD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                        ij,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -i- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GIYCOSTL HYDROLASES).
-i- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY J (FAMILY 44 OF GLYCOSTL HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Dacterium 'Caldocellum saccharolyticum',",
Appl. Environ. Microbiol. 57:694-700(1991).
-!- FUNCTION: DEGRADATION OF HEMICELLIOSES, THE SECOND MOST ABUNDANT
POLYSACCHARIDES IN NATURE. CONTAINS TWO CATALYTIC DOMAINS WITH
MANNANASE AND ENDOGLUCANASE ACTIVITIES.
-!- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-beta-D-mannosidic
                                                                                                                                                                                                                                                                                                                                                           01-JUL-1993 (Rel. 26, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
Beta-mannanase/endoglucanase A precursor [Includes: Mannan endo-1,4-beta-mannosidase A (EC 3.2.1.78) (Beta-mannase); Endo-1,4-beta-glucanase (EC 3.2.1.4) (Cellulase)].
                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                               4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -!- MISCELLANEOUS: THIS ENZYME IS MOST ACTIVE AT PH 6 AND 80 DEGREES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
Bacteria; Firmicutes; Bacillus/Clostridium group;
Thermoanaerobacter group; Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDIINE-91247819; Pubmed-2039230;
Luethi E., Jasmat N.B., Grayling R.A., Love D.R., Bergquist P.L.;
"Cloning, sequence analysis, and expression in Escherichia coli of
gene coding for a beta-mannanase from the extremely thermophilic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=93119139; PubMed=1476429; Gibbs M.D., Saul D.J., Luthi E., Bergquist P.L.; "The beta-mannanase from 'Caldocellum saccharolyticum' is part of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       galactoglucomannans. _
CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
                                                                                                        PROTON DONOR (BY SIMILARITY).
NUCLEOPHILE (BY SIMILARITY).
CELLULOSE-BINDING (BY SIMILARITY).
2E821E3D6BBACAO4 CRC64;
                                                                                                                                                                                                                                       Length 499;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    linkages in mannans, galactomannans, glucomannans, and
                                                                                                                                                                                                                                                                                        31; Indels
                                           Cellulose degradation; Hydrolase; Glycosidase; Signal.
                                                                                                                                                                                                                                         40.0%; Score 188; DB 1;
40.0%; Pred. No. 7.4e-14;
                                                                                                                                                                                                                                                                                     18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Appl. Environ. Microbiol. 58:3864-3867(1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1331 AA
                                                                                     ENDOGLUCANASE
CBD_3; 1.
GLYCOSYL_HYDROL_F5; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                              :||||: | ::: ||||:: 435
                                                                                                                                                                                                                                                                                                                                                                                                                     64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-AUG-1991 (Rel. 19, Created)
                                                                                                                                                                         55169 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 1-346 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linkages in cellulose.
                                                                                                                                                                                                                                                                                     34; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                              29
499
169
257
                                                                                                                                               350
499 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  multidomain enzyme.
                                                                                                                                                                                                                                                              Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FROM N.A.
ProDom; PD001947;
PROSITE; PS00659;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=44001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MANB_CALSA
P22533;
                                                                                                        ACT_SITE
ACT_SITE
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                            Query Match
                                                                SIGNAL
                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MANB_CALSA
                                                                                                                                                                                                                                                                                     Matches
ò
                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
```

```
modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             364 SGQIKVLYANKETNSTTNTIRPWLKVVNSGSSSIDLSRVTIRYWYTVDGERAQSAIS-DW 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
Bacteria; Firmicutes; Bacillus/Clostridium group;
Thermoanaerobacter group; Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 1516-1742 FROM N.A.
MEDLINE-91247819; PubMed-2039230;
Leuthi E., Jasmat N.B., Grayling R.A., Love D.R., Bergquist P.L.;
"Cloning, sequence analysis, and expression in Escherichia coli of gene coding for a beta-mannanase from the extremely thermophilic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-OCT-1996 (Rel. 34, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
Endoglucanase A precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase A)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BETA-MANNANSE/ENDOGLUCANASE A.
CATALYTIC (MANNANASE ACITVITY).
PRO/SER/THR-RICH (PT BOX).
SUBSTRATE-BINDING (POTENTIAL).
PRO/SER/THR-RICH (PT BOX).
CATALYTIC (ENDOGLUCANASE ACTIVITY).
PRO/SER/THR-RICH (PT BOX).
CATALYTIC (ENDOGLUCANASE ACTIVITY).
NUCLEOPHILE (BY SIMILARITY).
T -> P (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=95336703; PubMed=7612247;
Te'O V.S., Saul D.J., Bergquist P.L.;
"celA, another gene coding for a multidomain cellulase from the
extreme thermophile Caldocellum saccharolyticum.";
Appl. Microbiol. Biotechnol. 43:291-296(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 1; Length 1331;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         rPTPT -> RQHQHRQ (IN REF. 2).
FFBCA51BB8D8F0E0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                         Probom; PD001947; CBD_3; 2.
PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
Hydrolase; Glycosidase; Cellulose degradation; Slgnal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                39.7%; Score 186.5; DB 1;
42.5%; Pred. No. 3.5e-13;
live 18; Mismatches 31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1742 AA.
                                                                                                                                         EMBL, M36663; AAA72861.1; -.
PIR; B43745; B43745.
PIR; A49554, A48954.
HSSP; Q06861; INBC.
InterPro; IPR001956; CBD_3.
InterPro; IPR001957; Glyco_hydro_F5.
Pfam; PF001942; CBD_3; 2.
Pfam; PF00150; cellulase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TPTPTPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         423 AQIGASNVTFKFVKLSSSVSGADYYLE 449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62 AAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-AUG-1991 (Rel. 19, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            146892
                                                                                                                      EMBL; L01257; AAA71887.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Multifunctional enzyme
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity
nes 37; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-44001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (Cellulase A)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GUNA_CALSA
P22534;
```

ï

THE N-TERMINAL

```
01-JUL-1989 (Rel. 11, Last Sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2011 (Rel. 40, Last annotation update)
16-OCT-2011 (Rel. 40, Last annotation update)
17-OCT-2011 (Endo-1,4-betta-glucanase) (Cellulase)
(Cellobiohydrolase); Exoglucanase (EC 3.2.1.91) (Exocellobiohydrolase)
                                                                                                                                                                                                                                                                                                                                                                                                                                 reducing ends of the chains.
SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO CELLULASE FAMILY F(FAMILY 10 OF GLYCOSYL HYDROLASES).
SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                -i- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
linkages in cellulose.
-i- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glucosidic linkages
in cellulose and cellotetraose, releasing cellobiose from the non-
                                                                                                                                                                                                                              MEDLINE-89098398; PubMed=2789517;
Saul D.J., Williams L.C., LOVE D.R., Chamley I.W., Bergquist P.I.;
Saul D.J., Williams L.C., LOVE D.R., Caldocellum saccharolyticum encoding for exocellulase and endocellulase activity.";
Nuclect Cadas Res. 17:439-439(1989).
-i- FUNCTION: THIS PROPEIN IS MADE UP OF TWO DOMAINS: THE N-TERMINAL DOMAIN HAS EXOGLUCANASE ACTIVITY WHILE THE C-TERMINAL DOMAIN IS
                                                                                                                                 Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
Bacteria; Firmicutes; Bacillus/Clostridium group;
Thermoanaerobacter group; Caldicellulosiruptor.
                                                                                                    (1,4-beta-cellobiohydrolase)].
                                                                                                                                                                                                                                                                                                                                                AN ENDOGLUCANASE
                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                   NCBI_TaxID=44001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ij
                                                                                                                                                                                                                                                                                                               modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                             CATALTICE ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic linkages in cellulose.

PTW: THE LINKER REGION (ALSO TERMED "HINGE") MAY BE A POTENTIAL SITE FOR PROTEOLYSIS.

SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY E FRANTLY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY LIN THE C-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY L (FAMILY 48 OF GLYCOSYL HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C-TERMINAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                 Appl. Environ. Microbiol. 57:694-700(1991).
--i- FUNCTION: THE N-TERMINAL DOMAIN OF CELA ENCODES FOR AN
ENDOGLUCANASE ACTIVITY ON CARRENXETHICELLULOSE. THE C-TERMI
DOMAIN PROBABLY ACT SYNERGISTICALLY TO HYDROLYZE CRYSTALLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CELLULOSE-BINDING (BY SIMILARITY).
LIMKER ("HINGE"), (PRO-THR BOX).
CELLULOSE-BINDING (BY SIMILARITY).
LINKER ("HINGE") (PRO-THR BOX).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 1; Length 1742;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Signal; Repeat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CATALYTIC 1.
LINKER ("HINGE") (PRO-THR BOX)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3F0699A2123EED07 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; 1PR001556; CBD_3.
InterPro; 1PR001556; Glyco_hydro_48.
InterPro; 1PR001501; Glyco_hydro_9.
Pfam; PF00042; CBD_3; 3.
Pfam; PF0011; Glyco_hydro_49; 1.
Pfam; PF00759; Glyco_hydro_49; 1.
PRINTS; PR00844; GLHYDRIASE48.
Probom; PD001947; CBD_3; 2.
Probom; PD011903; Glyco_hydro_48; 1.
PROSITE; PS00592; GLYCOSYL_HYDROL_F9_2; 1.
PROSITE; PS00698; GLYCOSYL_HYDROL_F9_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                39.7%; Score 186.5; DB 1;
42.5%; Pred. No. 4.9e-13;
Live 18; Mismatches 31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ENDOGLUCANASE A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIMILARITY.
SIMILARITY.
SIMILARITY.
bacterium 'Caldocellum saccharolyticum'.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CATALYTIC 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     763 AQIGASNVTFKFVKLSSSVSGADYYLE 789
                                                                                                                                                                                                                                                                                                                                                                                               PIR; A43745; -; NOT_ANNOTATED_CDS. HSSP, P26221; 1TF4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ٨
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BY
                                                                                                                                                                                                                                                                                                                                                                               EMBL; L32742; AAA91086.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 193696
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1061
1113
396
434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                443
1545
1742
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Simi
Matches 37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     THR/PRO-RICH, TANDEM REPEATS OF T-P. CELLULOSE-BINDING (BY SIMILARITY).
THR/PRO-RICH, TANDEM REPEATS OF T-P. PROTON DONOR (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 1039;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ENDOGLUCANASE/EXOGLUCANASE B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NUCLEOPHILE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         31; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0E0378171594DDAE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                             Pfan: PF00331; Glyco_hydro_10; 1.

PRINTS; PR00134; GLHYDRLASE10.

PRODOM; P0001947; CBD_3; 1.

PROSITE; PS00591; GLYCOSYL_HYDROL_F10; 1.

Cellulose degradation; Hydrolase; Glycosidase; Repeat; Multifunctional enzyme; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     38.8%; Score 182.5; DB 141.9%; Pred. No. 7.5e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY
                                                                                                                                                  or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                       FIRSP, Q06811, INNC.
InterPro; IPR001956; CBD_3.
InterPro; IPR001000; Glyco_hydro_10.
InterPro: IPR001507; Glyco_hydro_F5.
Pfam; PF00192; CBD_3; 1.
Pfam; PF00150; cellulase; 1.
                                                                                                                                                                                              EMBL; X13602; CAA31936.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          117641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1039
                                                                                                                                                                                                                        PIR; S02711; S02711.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
Matches 36; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  177
285
792
1039
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qγ
```

1039 AA.

PRT;

STANDARD;

GUNB\_CALSA P10474;

GUNB\_CALSA

ŏ g

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-FEB-1995 (Rel. 31, Last sequence update)
01-NOV-1995 (Rel. 32, Last annotation update)
Endoglucanase I precursor (EC 3.2.1.4) (EGI) (Endo-1,4-beta-glucanase)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Gene sequence and properties of Cell, a family E endoglucanase from
                                                                                                                                            Hansen C.K., Diderichsen B., Joergensen P.L.; "celA from Bacillus lautus PL236 encodes a novel cellulose-binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 VSGCVKVQYKNND-SAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Clostridium thermocellum.
Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
Clostridium.
                                                                                                                                                                                                  endo-beta-1,4-glucanase.";
J. Bacteriol. 174:3522-3531(1992).
-!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIĀL.
ENDOGLUCANASE A.
CELLULOSE-BINDING (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 700;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN=NCIB 10682;
MEDLINE=93171873; PubMed=8436949;
Hazlewood G.P., Davidson K., Laurie J.I., Huskisson N.S.,
Gilbert H.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           27; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BY SIMILARITY.
3D5C8CADA53EEE0F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hydrolase; Glycosidase; Signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 172; DB 1;
Pred. No. 7.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    879 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      608 DWAKLGGSNIQISFGNHNGA--DSDTYAE 634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  60 DWAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A., AND SEQUENCE OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                      STRAIN=PL236;
MEDLINE=92276330; PubMed=1592807;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequ
01-NOV-1995 (Rel. 32, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              76910 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     36.6%;
44.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; M76588; AAA22303.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001956; CBD_3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00942; CBD_3; 1.
ProDom; PD001947; CBD_3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              40; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              700
700
213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Cellulose degradation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR; B41897; B41897.
PIR; S27499; S27499.
HSSP; Q06851; 1NBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 700 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                       HYDROLASES).
NCBI_TaxID=1401;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=1515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (Cellulase I).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GUNI_CLOTM
Q02934;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
SEQUENCE
                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GUNI_CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Dp,
  NAME OF THE PROPERTY OF THE PR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7 VQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMGC 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hansen C.K., Joergensen P.L., Diderichsen B.;
"celA from Bacillus lautus PL236 encodes a novel cellulose-binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15.7UL-1998 (Rel. 36, Last annotation update)
Endoglucanase A precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           э;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 145;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9514E3A71B106AEB CRC64;
                                                                                                                                                                                                                                                                                                             01-APR-1993 (Rel. 25, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
Hypothetical protein in CELA 5'region (Fragment).
Paenibacillus lautus (Bacillus lautus).
                                                                                                                                                                                                                                                                                                                                                                                                                               Bacteria; Firmicutes; Bacillus/Clostridium group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               38.4%; Score 180.5; DB 1;
43.9%; Pred. No. 1.2e-13;
iive 18; Mismatches 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Paenibacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacillus/Staphylococcus group; Paenibacillus
                                                                                                                                                                                                                                   145 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          700 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Paenibacillus lautus (Bacillus lautus).
                                                           :| |: : : || ||: || 479 QIGASNVTFKFVKLSSSSVSGADYYLE 504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                       63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          endo-beta-1,4-glucanase.";
J. Bacteriol. 174:3522-3531(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE=92276330; PubMed=1592807;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE 145 AA; 15782 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                60 SNIQISFG--NHTGTNSDTYVE 79
                                                                                                                                                                                                                                                                                     (Rel. 25, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; M76588; AAA22302.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            67 GNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HSSP; Q06851; INBC.
InterPro; IPR001956; CBD_3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FIAM; FF00942; CBD_3; 1.
ProDom; PD001947; CBD_3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF00942; CBD_3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (Cellulase A) (EG-A).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PIR; A41897; A41897.
PIR; S27498; S27498.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      pothetical protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 36; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=1401;
                                                                                                                                                                                                                                                                                     01-APR-1993
                                                                                                                                                                                                                                CEA_PAELA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GUNA_PAELA
```

GUNA\_PAELA

ă g ò qq CELA.

., M

Gaps

4

```
CIPA_CLOTM
Q06851;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REPEAT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                             NON_TER
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CIPA_CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ID DIT DIT OCC OCC OCC OCC
 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Op
                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 : | | : | | | : | | | : | | | 36 IKGEVVLQYANGNAGATSNSINPRFKIINNGTKAINLSDVKIRYYYTKEGGASQNFW-CD 794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-0CT-1994 (Rel. 30, Created)
01-0CT-1994 (Rel. 30, Last sequence update)
16-0CT-2001 (Rel. 40, Last annotation update)
Cellulosomal scaffolding protein B (Cellulosomal glycoprotein S1/SL)
(Cellulose integrating protein B) (Fragment).
       J. Gen. MICROBIOL. 139:307-316(1993).

-1- FUNCTION: THIS BRIXYME CATALYZES THE ENDOHYDROLYSIS OF 1.4-BETA-GLUCOSIDIC LINKAGES IN CELLULOSE, LICHENIN AND CEREAL BETA-D-GLUCANS. PRINCIPALLY ACTIVE AGAINST BARLEY BETA-GLUCAN.

-1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic linkages in cellulose.

-1- PATHWAY: CELLULOSE DEGRADATION.

-1- SIMILARITY: BELONGS TO CELLULASE FAMILY E (FAMILY 9 OF GLYCOSYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Clostridium thermocellum.
Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
Clostridium.
                                                                                                                                                                                                                                                                                                                                                                            CATALYTIC.
CELLULOSE-BINDING (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   35.2%; Score 165.5; DB 1; Length 879; 33.0%; Pred. No. 5.3e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=93146373; PubMed=1490597;
Poole D.M., Morag E., Lamed R., Bayer E.A., Hazlewood G.P.,
Gilbert H.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           32; Indels
                                                                                                                                                                                                                                                                                                                                                                                               SIMILARITY.
( SIMILARITY.
( SIMILARITY.
35A60069A514A927 CRC64;
                                                                                                                                                                                                                                                                                                                        Cellulose degradation; Hydrolase; Glycosidase; Signal. Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                26; Mismatches 32
                                                                                                                                                                                                                              PIR; A47704; A47704.

PIRSP; A47704; A47704.

HSSP; P26221; 1FF4.

InterPro; IPR001956; CBD_3.

InterPro; IPR001701; Glyco_hydro_9.

Pfam; PF00782; CBD_3; 2.

Pfam; PF00789; Glyco_hydro_9; 1.

PROSITE; PS001947; CBD_3; 1.

PROSITE; PS00698; GLYCOSYL_HYDROL_F9_1; 1.

PROSITE; PS00698; GLYCOSYL_HYDROL_F9_2; 1.
                                                                                                                                                                                                                                                                                                                                                               ENDOGLUCANASE I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 772 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                  97796 MW;
                                                                                                                                                                                                                      EMBL; L04735; AAA20892.1; -.
Clostridium thermocellum.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                              56
56
729
448
486
495
879 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                      HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=1515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CIPB_CLOTM
Q01866;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           29;
                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
ACT_SITE
ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local
Matches 2
                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CIPB_CLOTM
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
Indels 10; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae; Clostridium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-NOV-1997 (Rel. 35, Last sequence update)
1-MAR-2002 (Rel. 41, Last annotation update)
Cellulosomal scaffolding protein A precursor (Cellulosomal glycoprotein S1/SL) (Cellulose integrating protein A) (Cohesin).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            X 147 AA APPROXIMATE REPEATS. (PARTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           35.1%; Score 165; DB 1; Length 772; 36.1%; Pred. No. 5.2e-11; Live 23; Mismatches 29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 X 24 AA APPROXIMATE REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Probom; PD001947; CBD_3; 1.
PROSTIE: PS00018: EF_HAND; UNKNOWN_1.
PROSTIE: PS000448: CLOS_CELLUJOSOME RPT; 2.
Cellulose degradation; Cell wall; Glycoprotein; Repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BBF06DE5E094FE10 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRO/THR-RICH.
CELLULOSE-BINDING.
PRO/THR-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     336 HAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLE 372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 1853 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 WAAM ------GCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRO/THR-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001956; CBD_3.
InterPro; IPR002102; Cohesin.
InterPro; IPR002105; Dockerin_1.
InterPro; IPR002048; EF-hand.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-JUN-1994 (Rel. 29, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF00942; CBD_3; 1.
Pfam; PF00963; Cohesin; 3.
Pfam; PF00404; Dockerin_1; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        82491 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; X68233; CAA48312.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 35.1%
Best Local Similarity 36.1%
Matches 35; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Clostridium thermocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 710
710
743
772 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               462
462
81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   006851;
```

us-09-917-376-4.rsp

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        between the Swiss Institute of Bioinformatics and the EMBL outstation—the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                       MEDLINE-93209931; PubMed-8458832; Fujino T., Beguin P., Aubert J.-P.; "Organization of a Clostridium thermocellum gene cluster encoding the cellulosomal scaffolding protein CipA and a protein possibly involved In attachment of the cellulosome to the cell surface."; J. Bacteriol. 175:1891-1899(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CELLULOLYTIC ENZYMES.
--- SUBCELLULAR LOGATION: CELL SURFACE.
--- DOMAIN: CONTAINS 9 COPIES OF A DOMAIN (COHESIN) OF ABOUT 150
RESIDUES THAT BINDS TO THE DOCKERIN DOMAIN BORN BY THE CATALYTIC
--- COMPONENTS OF THE CELLULOSOME.
--- DOMAIN: A 24 RESIDUES DOMAIN IS REPEATED TWICE IN THIS ENZYME AS WELL AS IN OTHER C.THERMOCELLUM CELLULOSOME ENXYMES. THIS DOMAIN MAY FUNCTION AS THE BINDING LIGAND FOR THE SL COMPONENT.
                                                                                                                                                                                                                                                                                                                                                                                  Shimon L.J.W., Bayer E.A., Morag E., Lamed R., Yaron S., Shoham Y.,
                                                                                                                   "Sequencing of a Clostridium thermocellum gene (cipA) encoding the cellulosomal SL-protein reveals an unusual degree of internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-97076134; Pubmed-8918451;
Tormo J., Lamed R., Chirino A.J., Morag E., Bayer E.A., Shoham Y.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Tavares G.A., Beguin P., Alzari P.M.;
"The crystal structure of a type I cohesin domain at 1.7-A
resolution.";
J. Mol. Biol. 273:701-713(1997).
-i- FUNCTION: ACTS AS A SCAFFOLDING PROTEIN IN THE CELLULOSOME. IT
PROMOTES BINDING OF CELLULOSE TO THE CATALYTIC DOMAINS OF THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Crystal structure of a bacterial family-III cellulose-binding
                                                                                      Serngross U.T., Romaniec M.P.M., Kobayashi T., Huskisson N.S.,
                                                                                                                                                                                                                                                                                                                                                                                                               A cohesin domain from Clostridium thermocellum: the crystal
structure provides new insights into cellulosome assembly.";
Structure 5:381-390(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             domain: a general mechanism for attachment to cellulose."
EMBO J. 15:5739-5751(1996).
                                  SEQUENCE FROM N.A., AND SEQUENCE OF 29-40 AND 42-43.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 1213-1361
MEDLINE=98022914; PubMed=9402065;
                                                                                                                                                                                                                                                                                                                                                X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 184-321.
MEDLINE-97238934; PubMed-9083107;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 368-522.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; L08665; -; NOT_ANNOTATED_CDS.
EMBL; X67506; CAA47840.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR001956; CBD_3.
InterPro; IPR002102; Cohesin.
InterPro; IPR002105; Dockerin_1.
InterPro; IPR002048; EF-hand.
Pfam; PF00942; CBD_3; 1.
Pfam; PF00963; Cohesin; 9.
                                                  STRAIN-ATCC 27405;
MEDLINE-93302508; PubMed-8316083;
                                                                                                                                                                        Mol. Microbiol. 8:325-334(1993)
                                                                                                                                                                                                          SEQUENCE OF 1820-1853 FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PIR; S36859; S36859.
PDB; IANU; 23-JUL-97.
PDB; IAOH; 08-JUL-98.
PDB; INBC; 26-SEP-97.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Steitz T.A.;
                                                                                                                                                                                                                                                                                                                                                                                                   Frolow F
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                             5
                                                                                                    CELLULOSOMAL SCAFFOLDING PROTEIN A.
CELLULOSE-BINDING (BY SIMILARITY).
9 X 150 AA APPROXIMATE TANDEM REPEATS.
1-1.
1-2.
1-3.
1-4.
1-5.
1-6.
1-7.
1-8.
2 X 24 AA APPROXIMATE REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
STRAIN-ATROSEPTICA FCBR C18;
MEDLINE-9829944; Pubmed-9636315;
Olsen O., Thomsen K.K., Weber J., Duus J.O., Svendsen I., Wegener C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ||| :||:: |:: : | | | | :: ||||| :||| :||| :||| 365 VSGNLKVEFYNSNPSDTTNSINPQFKVTNTGSSAIDLSKLTLRYYYTVDGQKDQTFW-CD 423
                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 linkages in cellulose.
-!- SUBCELLULAR LOCATION: Secreted.
-!- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLXCOSYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 von Wettstein D.;
"Transplanting two unique beta-glucanase catalytic activities into multiensyme, which forms glucose.";
Biotechnology 14:71-76(1996).
-:- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1997 (Rel. 35, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
Endoglucanase N precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase N)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
           Prodom; P0001947; CBD_3; 1.
PROSITE; PS00018; EF_HAND; UNKNOWN_1.
PROSITE; PS00448; CLOS_CELLULOSOME_RPT; 2.
Cellulose degradation; Cell wall; Glycoprotein; Repeat; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                             35.1%; Score 165; DB 1; Length 1853; 36.1%; Pred. No. 1.5e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                              -> AA (IN REF. 1).
3ABDDC03ABFC5372 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    444 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 WAAM------GCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                            23; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                   Α.
Μ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequ
01-NOV-1997 (Rel. 35, Last anno
                                                                                                                                                                                                                                                                                                                                                                196831
 Dockerin_1;
                                                                                                                                                                                                                                                                                                                                                                                                                                             35; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                      182
322
704
704
866
1031
1196
1361
1526
                                                                                                                                                                                                                                                                                                                                                                  AA;
                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Erwinia carotovora
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HYDROLASES).
                                                                                                                                                                                                                                                                   1384
1548
1791
1791
1824
1615
                                                                                                                                                                                      560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pectobacterium.
NCBI_TaxID=554;
PF00404;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (Cellulase N)
                                                                             3D-structure
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GUNN_ERWCA
Q59394:
                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                CONFLICT
                                                                                                          CHAIN
                                                                                                                                                                                                                     REPEAT
REPEAT
REPEAT
                                                                                                                                                        REPEAT
REPEAT
                                                                                                                                                                                      REPEAT
REPEAT
                                                                                                                                                                                                                                                                   REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                  REPEAT
                                                                                             SIGNAL
                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GUNN_ERWCA
                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
```

œ

```
ò
                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                             5
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOI. Gen. Genet. 241:341-350(1993).
-!- FUNCTION: ENDOGLUCANASE WITH SOME EXOGLUCANASE ACTIVITY. THE PH
OPTIMUM IS ABOUT 7.0 AND THE TEMPERATURE OPTIMUM ABOUT 42 DEGREES
                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                  SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDG--GSSTLVYNC 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linkages in cellulose.
SUBCELLULAR LOCATION: Secreted.
SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Cooper V.J.C., Salmond G.P.C.; "Molecular analysis of the major cellulase (CelV) of Erwinia carotovora: evidence for an evolutionary 'mix-and-match' of enzyme
                                                                                                                                                                                                                                                                                                                                                                                                                01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
Endoglucanase V precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Pectobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
                                                                                                                                                                                                                                             .,
.,
                                                                                                                                                                                                                      DB 1; Length 444;
                                                                                                                                                           PROTON DONOR (BY SIMILARITY).
NUCLEOPHILE (BY SIMILARITY).
FA7E4179004CBB43 CRC64;
                                                                                                                                                                                                                   31.8%; Score 149.5; DB 1; Length 38.6%; Pred. No. 1.6e-09; Live 19; Mismatches 30; Indels
                                           HSSP; 085465; 1A3n.
InterPro; IPR001956; CBD_3.
InterPro; IPR0019547; Glyco_hydro_F5.
Pfam; PF00942; CBD_3; 1.
Pfam; PF00150; cellulase; 1.
ProDom; PF001947; CBD_3; 3.
PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
31 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                             505 AA.
                                                                                                                                                                                                                                                                                                                            87
                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                               DWAAMGCGNIRASFGSVNPATPTADTYL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=94067016; PubMed=8246888;
                                                                                                                                                                                   48300 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; X76000; CAA53592.1; -.
                                EMBL; L39788; AAC37033.1; -.
                                                                                                                                                                                                            Query Match
Best Local Similarity 38.69
                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                           168
256
                                                                                                                                                           168
256
444 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Erwinia carotovora
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=554;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HYDROLASES)
                                                                                                                                                                                                                                                                                                                                                                                                                                                               (Cellulase V).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-SCRI193
                                                                                                                                                                                                                                                                                                                                                                                           GUNV_ERWCA
Q47096;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   domains.";
                                                                                                                                                           ACT_SITE
ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                  GUNV_ERWCA
ò
                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                               ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDG--GSSTLVYNC 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
Endoglucanase VI precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase VI)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mae A., Helkinhelmo R., Palva E.T.;
"Structure and regulation of the Erwinia carotovora subspecies carotovora SCC3193 cellulase gene celv1 and the role of cellulase in phytopathogenicity.";
Mol. Gen. Genet. 247:17-26(1995).
--- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            linkages in cellulose.
-:- SUBCELLULAR LOCATION: Secreted.
-:- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              .,
.,
                                                                                                                                                                                                                                                                                                          (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 505;
                                                                                                                                                                                                                                                                                                                            PROTON DONOR (BY SIMILARITY).
NUCLEOPHILE (BY SIMILARITY).
DBEA9337BB4D2623 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                     Probom; PD001947; CBD_3; 1.
PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
Cellulose degradation; Hydrolase; Glycosidase; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       31.8%; Score 149.5; DB 1; 38.6%; Pred. No. 1.8e-09; tive 19; Mismatches 30;
                                                                                                                                                                                                                                                                             LINKER.
CELLULOSE-BINDING
                                                                                                                                                                                                                         ENDOGLUCANASE V. CATALYTIC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        504 AA.
                                                                                                                                                                                                      POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              60 DWAAMGCGNIRASFGSVNPATPTADTYL 87
                    InterPro; IPR001956; CBD_3.
InterPro; IPR001547; Glyco_hydro_F5.
Pfam; PF00942; CBD_3; 1.
Pfam; PF00150; cellulase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HSSP; O85465; IA3H.
INTERPROJ 1PR001956; CBD_3.
InterPro; IPR001547; Glyco_hydro_F5.
Pfam; PF00942; CBD_3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=95231512; PubMed=7715600;
                                                                                                                                                                                                                                                                                                                                                                                   54900 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; X79241; CAA55823.1; -. HSSP; O85465; 1A3H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                               353
168
256
505 AA;
HSSP; 085465; 1A3H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
Matches 34; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Erwinia carotovora.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pectobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=554;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (Cellulase V1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-SCC3193;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GUNW_ERWCA
Q59395;
                                                                                                                                                                                                                                                                                                                            ACT_SITE
ACT_SITE
SEQUENCE
                                                                                                                                                                                                                           CHAIN
DOMAIN
                                                                                                                                                                                                      SIGNAL
                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GUNW_ERWCA
```

```
g
                   ò
                                                                                                                                                                                                                                                                                                            δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                   ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDINE-91364686; PubMed-1999625; Bronnenmeier K., Ruecknagel K.P., Staudenbauer W.L.; Bronnenmeier K., Ruecknagel K.P., Staudenbauer W.L.; Bronnenmeier M.D.; Purlitation and properties of a novel type of exo-1,4-beta-glucanase (avicelase II) from the cellulolytic thermophile Clostridium stercorarium.".

Eur. J. Biochem. 200:379-385(1991).
-!- CATALTYTE ACTIVITY: Hydrolysis of 1,4-beta-D-glucosidic linkages in cellulose and cellotetraose, releasing cellobiose from the non-reducing ends of the chains.
-!- SIMILARITY: BELONGS TO CELLULASE FAMILY L (FAMILY 48 OF GLYCOSYL
                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                               2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDG--GSSTLVYNC 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
Exoglucanase II precursor (EC 3.2.1.91) (Exocellobiohydrolase II)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Clostridium stercorarium.
Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
Clostridium.
                                                                                                                                                                                                                                                   5;
                                                                                                                                  CELLULOSE-BINDING (BY SIMILARITY)
                                                                                                                                                                                                                      Length 504;
                                                                                                                                             PROTON DONOR (BY SIMILARITY).
NUCLEOPHILE (BY SIMILARITY).
OD7ECF74781565FA CRC64;
                                                                                                                                                                                                                                                   31; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Staugenbauer w.L.;
Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
Pfam; PF00150; cellulase; 1.
Probom; PD001947; CBD_3; 1.
PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
Cellulose degradation; Hydrolase; Glycosidase; Signal.
31 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bronnenmeier K., Kundt K., Riedel K., Schwarz W.H., Staudenbauer W.L.;
                                                                                                                                                                                                                     Query Match

31.4%; Score 147.5; DB 1;
Best Local Similarity 37.5%; Pred. No. 3.1e-09;
Matches 33; Conservative 19; Mismatches 31;
                                                                                      ENDOGLUCANASE V1
CATALYTIC.
                                                                                                                                                                                                                                                                                                                                                                                                                                            914 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                        60 DWAAMGCGNIRASFGSVNPATPTADTYL 87
                                                                                                                     LINKER
                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                            54963 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; 269359; CAA93280.1; -. HSSP; Q06851; 1NBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                            31
504
334
352
504
168
                                                                                                                                 353
168
256
504 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN=NCIB 11754;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN=NCIB 11754;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHARACTERIZATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-1510;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HYDROLASES)
                                                                                                                                                                                                                                                                                                                                                                                                                                             GUX2_CLOSR
                                                                                                                                             ACT_SITE
ACT_SITE
SEQUENCE
                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                        P50900;
                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                            GUX2_CLOSR
                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
                  SORFIFFFFFF
                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                  g
```

InterPro; IPR001956; CBD\_3

```
9
                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                    1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                      31.4%; Score 147.5; DB 1; Length 914; 35.2%; Pred. No. 6.2e-09; tive 21; Mismatches 35; Indels 1.
                                                                                                                                                                              EXOGLUCANASE II.
7; DODB6017D6DFF82C CRC64;
InterPro; IPR000556; Glyco_hydro_48.

Pfam; PF00942; CBD_3; 1.

Pfam; PF02011; Glyco_hydro_48; 1.

PRINTS; PR00844; GLHYDRLASE48.

ProDom; PD011903; Glyco_hydro_48; 1.

ProDom; PD011903; Glyco_hydrolase; Glycosidase; Signal.

SIGNAL 1 33 POTENTAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Search completed: July 2, 2002, 09:18:04 Job time: 372 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 WAAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                CHAIN 34 914 EX
SEQUENCE 914 AA; 103020 MW;
                                                                                                                                                                                                                                                                                           Best Local Similarity 35.23
Matches 31; Conservative
                                                                                                                                                                                                                                                                           Query Match
```

1;

083012 bacillus sp 04510 bacillus sp 09510 acetivibrio 091138 clostridium 091039 bacteroides 097774 clostridium 04596 clostridium 08230 clostridium 082319 bacitlus sp 046392 clostridium 04532 clostridium 04532 clostridium 04532 clostridium 04532 phono sapien 09542 ginglymosto 09542 ginglymosto 09542 ginglymosto 09932 caenorhabdi 09333 autonomous penos autonomous

09y775 candida tro 09yhp9 ginglymosto 098hki rhizobium 1 093073 hepatiis g

1;

Gaps

1;

AMGCGNIRASFGSVNPATPTADTYLQ 88

63

ò

bacillus bacillus

Q93TJ6 Q93LD0

190 187 186

55 10 10 10 10 10 10 10 10

. 9

Result

QQ

GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA

```
01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAY-2000 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
HYPOTHETICAL 18.5 KDA PROTEIN (FRAGMENT).
Caldibacillus cellulovorans.
Bacteria, Firmicutes; Bacillus/Clostridium group;
Bacillus/staphylococcus group; Alicyclobacillus group; Caldibacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-20120520; PubMed-10653733; MEDLINE-20120520; PubMed-10653733; Sunna A., Glbbs M.D., Chin C.W.J., Nelson P.J., Bergquist P.L.; Sunna A., Glbbs M.D., Chin C.W.J., Nelson P.J., Bergquist P.L.; Agene encoding a novel multidomain beta-1.4-mannanase from Caldibacillus cellulovorans and action of the recombinant enzyme on kraft pulp."; Appl. Environ. Microbiol. 66:664-670(2000). EMBL; AF163837; AAF22273.1; ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 2; Length 170;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
48.0%; Score 225.5; DB 2; Length
Best Local Similarity 47.7%; Pred. No. 2.6e-17;
Matches 41; Conservative 17; Mismatches 27; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hypothetical protein.

NON_TER 1 1 1 SEQUENCE 170 AA; 18493 MW; 7AC9D33F44E3A0B4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                              ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  170
0083012
0087610
00976130
0097614
0097714
0097714
00973130
0097312
0097312
0097312
0097312
0097312
0097312
0097312
0097312
0097312
0097312
0097312
0097312
0097312
0097312
0097312
0097312
0097312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                13
                                                                                                                                                                                                                                  115
115
117
118
119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR001956; CBD_3. Pfam; PF00942; CBD_3; 1. ProDom; PD001947; CBD_3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
501
19186
4733
19186
1115483
112483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
1115483
111548
111548
111548
111548
111548
111548
111548
111548
111548
11548
11548
11548
11548
11548
11548
11548
11548
11548
11548
11548
1
Q9RFX6
 RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9RFX6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2, 2002, 09:17:33 ; Search time 91.42 Seconds (without alignments) 168.416 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    anaerocellu
caldicellul
caldicellul
bacillus su
bacillus su
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   bacillus sp
caldicellul
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9rfx6 caldibacill Q9rfx5 caldibacill
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              thermophili
caldicellul
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          caldicellul caldicellul
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         caldibacill
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Description
                                                                                                                                                                            1 VSGGVKVQYKNNDSAPGDNQ......RASFGSVNPATPTADTYLQX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      091818
092411
093418
0934310
093395
09330
093316
093116
093116
093116
                           Compugen Ltd
                                                                                                                                                                                                                                                                        Total number of hits satisfying chosen parameters:
            GenCore version 4.5
Copyright (c) 1993 - 2000 Comp
                                                                                                                                                                                                                                              562222 seqs, 172994929 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                        Listing first 45 summaries
                                                                   OM protein - protein search, using sw model
                                                                                                                                                                                                                  Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9RFX5
Q9L8L8
Q9Z4I1
Q9AQG4
Q9X3P5
Q9AX3P6
Q9AX3P6
Q9AX3P6
Q9AX3P6
Q9AQH0
Q9AQH0
Q9AQH0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_rodent:*
sp_virus:*
sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9RFX6
                                                                                                                                                                                                                                                                                                                                                                                                                                                       sp_human:*
sp_invertebrate:*
sp_mammal:*
sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sp_rvirus:*
sp_bacteriap:*
sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                            Match 100%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_organelle:*
sp_phage:*
sp_plant:*
                                                                                                                                                                                                                                                                                                                                                                                                             sp_archea:*
sp_bacteria:*
sp_fungi:*
                                                                                                                                                                                                                                                                                                    Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                            Post-processing: Minimum Match 0%
                                                                                                                                               US-09-917-376-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB
                                                                                                                                                                                                                                                                                                                                                                                                   SPTREMBL_19:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1751
1000
1770
1770
1711
1996
1779
499
499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length
                                                                                                                                                                                                       BLOSUM62
                                                                                                                                                                                                                                                                                                                                                            Maximum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          July
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                        225.5
216.5
214.5
204.5
194.5
192.5
192.5
191.5
191.5
191.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score
                                                                                                                                                               Perfect score:
                                                                                                                                                                                                       Scoring table:
```

..

Database

Searched:

Sequence:

Title:

ë O

Run

RESULT Q9RFX5

A PART REPORT OF THE PROPERTY 
g

```
STRAIN=BP-23; MEDINE-21129642; PubMed-11234960; MEDINE-21129642; PubMed-11234960; MEDINE-21129642; PubMed-11234960; Pastor F.I.J., Pujol X., Blanco A., Vidal T., Torres A.L., Diaz P.; Molecular cloning and characterization of a multidomain endoglucanase from Paenibacillus sp BP-23: evaluation of its performance in pulp refining.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "A novel thermostable multidomain 1,4-beta-xylanase from 'Caldibacillus cellulovorans' and effect of its xylan-binding domain on enzyme activity."

Microbiology 146:2947-2958(2000).

EMBL; AF200304; AAF61649.1; -.
HSSP; Q06861; NNBC.

InterPro; IPR001365; CBD_3.

InterPro; IPR001305; CBD_6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 L APPL. Microbiol. Biotechnol. 55:61-68(2001).

1 - SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.

R EMBL, AJ133614; CAB38941.1; --
R HSSP: P26221; 1TF4.

R InterPro; IPR001956; CBD_3.

R InterPro; IPR001956; CBD_3.

R InterPro; IPR001961; FN_III.

R InterPro; IPR001701; Glyco_hydro_9.

R Pfam; PF00042; CBD_3; 2.

R Pfam; PF00759; Glyco_hydro_9; 1.

R ProDom; PD001941; CBD_3; 2.

R ProDom; P0001941; CBD_3; 2.

R ProDom; P0001941; CBD_3; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           45.6%; Score 214.5; DB 2; Length 921; 46.5%; Pred. No. 3.3e-15; tive 16; Mismatches 29; Indels 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL.
BETA-1,4-XYLANASE XYNA.
N; C5DDD1A7F7567413 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-MAY-1999 (TrEMBLrel. 10, Created)
01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
CELLULASE PRECURSOR (EC 3.2.1.4).
                                                                                                                                                                                                                                                                                                            InterPro; IRR001000; Glyco_hydro_10.
InterPro; IRR002965; P_rich_extensn.
Pfam; PF02018; CBD_6; 1.
Pfam; PF00311; Glyco_hydro_10; 1.
PRINTS; PR00134; GLHYDRIASE10.
PRINTS; PR01217; PRICHEXTENSN.
ProDom; PD001947; CBD_3; 2.
Signal; Xylan degradation; Hydrolase; Glycosidase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
NCBI_TaxID=89769;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            997 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PROSITE; PSO0592; GLYCOSYL_HYDROL_F9_1; 1. PROSITE; PSO0698; GLYCOSYL_HYDROL_F9_2; 1. PROSITE; PSO0213; LIPOCALIN; UNKNOWN_1. Glycosidase; Hydrolase; Signal. 34 SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          830 WVGCSNLRGSLVKLTTGRTGADYYLE 855
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                921 AA; 102380 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best_Local Similarity 46.5%
Matches 40; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           921
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BP-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Bacillus sp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q924I1
Q924I1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9Z4I1
                RRT RET DE BOR D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE—20120520; PubMed=10653733;
MEDLINE—20120520; PubMed=10653733;
MASUNDA A., Glbbs M.D., Clin C.W.J., Nelson P.J., Bergquist P.L.;
MA gene encoding a novel multidomain beta-1,4-mannanase from caldibacillus cellulovorans and action of the recombinant enzyme on traft pulp."

MA plub."

Appl. Environ. Microbiol. 66:664-670(2000).

BMBL; AF163837; AAF22274.1;

MA PROBEST, INNE.

InterPro; IPR004302; Chlin_bind_3.

InterPro; IPR004302; Chlin_bind_3.

InterPro; IPR004295; Prich_extensn.

Pfam; PF0042; CBD_3; 2.

Pfam; PF0042; CBD_3; 2.

Pfam; PF0042; CBD_3; 2.

Pfam; PF0150; cellulase; 1.

PR Pfam; PF0150; cellulase; 1.

PR Pfam; PF01217; PRICHEXTENSN.

PRODOM: P0001947; CBD_3; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Alicyclobacillus group; Caldibacillus.
NCBL_TaxID=74586;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Alicyclobacillus group; Caldibacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 46.1%; Score 216.5; DB 2; Length 930; Best Local Similarity 46.5%; Pred. No. 2e-15; Matches 40; Conservative 17; Mismatches 28; Indels 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 33 POTENTIAL.
34 930 MULTIDOMAIN BETA-1,4-MANNANASE.
930 AA; 101576 MW; 0086638D54D1A2CC CRC64;
                                                                                                                                                                                                                                                                                                                                                          01-MXY-2000 (TrEMBLrel. 13, Last sequence update) 01-DEC-2001 (TrEMBLrel. 19, Last annotation update) MULTICOMAIN BETA-1,4-MANNANSE PRECURSOR. MANA. Caldibacillus cellulovorans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q9L8L8;
01-CCT-2000 (TrEMBLrel. 15, Created)
01-CCT-2000 (TrEMBLrel. 15, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
BETA-1,4-XYLANASE XYNA PRECURSOR.
                                                                                                                                                                                                                                                   930 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    921 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A. MEDLINE=20519260; PubMed=11065373; Sunna A., Glbbs M.D., Bergquist P.L.;
                                                                                                                                                                                                                                                                                                                             -MAY-2000 (TrEMBLrel. 13, Created)
            79 QVGCSNVRGSFVKLSTGRTGADXYIE 104
                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  :|| |:| |: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |:: || |: || |: || |: || |: || |: || |:  || |:  || |:  || |:  || |:  || |:  || |:  || 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Caldibacillus cellulovorans
                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=74586;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIGNAL
                                                                                                                                                                                                                                              O9RFX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            091818
                                                                                                                                                                    N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      m
```

XYNA.

RESULT Q9L8L8

g à

ò

ï

Gaps

;;

us-09-917-376-4.rspt

g

ò a ò g

```
Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P., Bergquist P.L.;
"Multidomain and multifunctional glycosyl hydrolases from the extreme thermophile Caldicellulosiruptor isolate Tok7B.1.";
Curr. Microbiol. 40:333-340(2000).
EMBL, AF078737; AAD30363.1;
HSSP: Q06851: 1NBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                                                                                                Miyake K., Machida Y., Hattori K., Iijima S.; "Characterization of a multi-domain cellulase from an extremely thermophilic anaerobe strain NA10."; submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases. EMBL, AB008029; BAAA22939.1; -- HSSP; Q06851; INBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 2; Length 1000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             A RICEPTOR I PRO01956; CBD_3.

A InterPro: IPR001956; CBD_3.

A InterPro: IPR001956; CBD_3.

A InterPro: IPR001000; Glyco_hydro_10.

B Ffam; PF00942; CBD_3; 1.

A Ffam; PF00942; Cellulase; 1.

A Ffam; PF00150; Cellulase; 1.

A Ffam; PF00134; Glyco_hydro_10; 1.

A PRINTS; PR00134; CBD_3; 1.

A PROSITE; PS001947; CBD_3; 1.

A PROSITE; PS00591; GLYCOSYL_HYDROL_F10; 1.

A PROSITE; PS00591; GLYCOSYL_HYDROL_F5; UNKNOWN_1.

A PROSITE; PS00591; GLYCOSYL_HYDROL_F5; UNKNOWN_1.

A PROSITE; PS00591; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  41.0%; Score 192.5; DB 2; Length 43.7%; Pred. No. 9.9e-13; Live 17; Mismatches 31; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Thermoanaerobacter group; Caldicellulosiruptor. NCBI_TaxID=80339;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 1770 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001956; CBD_3.
InterPro; IPR001305; CBD_6.
InterPro; IPR001000; Glyco_hydro_10.
InterPro; IPR002965; P_rich_extensn.
Pfam; PF00942; CBD_3; 3.
Pfam; PF02018; CBD_6; 2.
Pfam; PF020131; Glyco_hydro_10; 1.
PRIWRS; PR00134; GLHYDRAASE10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 431 AQIGASNVTFKFVKLSSSVSGADYYLE 457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN=TOK7B.1;
MEDLINE=20171169; Pubmed=10706665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (TrEMBLrel. 12, C
(TrEMBLrel. 12, I
(TrEMBLrel. 19, I
                                      thermophilic anaerobe NA10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
es 38; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                      NCBI_TaxID=67756;
   BETA-GLUCANASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1999 (
01-NOV-1999 (
01-DEC-2001 (
                                                                                                                                                                                                                         STRAIN-NA10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09X3P5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DDT TO DO THE LEAD TO THE LEAD TO THE LEAD TO THE LEAD TH
   DER RELEASE DER RE
                                                                                                                                                                                      ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-TOK7B.1;
MEDLINB-20171169; PubMed=10706665;
Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P.,
Bergquist P.L.;
"Multidomain and multifunctional glycosyl hydrolases from the extreme
thermophile caldicellulosiruptor isolate Tok7B.1.";
Curr. Microbiol. 40:333-340(2000).
EMBL; AF076042; AAK06394.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                       2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 1751;
                                                                                                          Length 997;
                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PROSITE; PS00592; GLYCOSYL_HYDROL_F9_1; 1.
PROSITE; PS00698; GLYCOSYL_HYDROL_F9_2; 1.
PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
SEQUENCE 1751 AA; 192176 MW; 60178CBF3C00BE95 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
F20CB119D6410629 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Caldicellulosiruptor sp. Tok7B.1.
Bacteria; Firmicutes; Bacillus/Clostridium group; Thermoanaerobacter group; Caldicellulosiruptor.
                                                                                                   Query Match
43.5%; Score 204.5; DB 2;
Best Local Similarity 43.7%; Pred. No. 4.6e-14;
Matches 38; Conservative 21; Mismatches 27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          41.4%; Score 194.5; DB 2; 44.7%; Pred. No. 1.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 44.7%; Pred. No. 1.2e-
Matches 38; Conservative 14; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 1000 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 1751 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HSSP, 008851; 1NBC.
InterPro; IPR001956; CBD_3.
InterPro; IPR001956; CBD_3.
InterPro; IPR001965; Lipocln_cytFABP.
InterPro; IPR002965; P_rich_extensn.
Pfam; PF00945; CBD_3; 1.
Pfam; PF00759; Glyco_hydro_9; 1.
PRIWRS; PR01217; PRICHEXTENSN.
ProDom; PD001947; CBD_3; 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                              | :| |: | :|||| ||||||: | 905 AQVGSANVQGMFVAVNPAKGTADTYVE 931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Created)
                                                                                                                                                                                                                                                                                                                                                                                                        88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                    62 AAMGCGNIRASFGSVNPATPTADTYLQ
106927 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            05,
05,
19,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      O24820;
01-JAN-1998 (TrEMBLrel.
01-JAN-1998 (TrEMBLrel.
01-DEC-2001 (TrEMBLrel.
   AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
997
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    024820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                09A0G4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              024820
```

g

ò g δ 12444 1444

Ĥ

4

```
Caldicellulosiruptor sp. Tok7B.1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Zverlov V.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           P96311;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P96311
                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 10
   à
                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;;
                                                                                                                Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P., Bergquist P.L.,
"Multidomain and multifunctional glycosyl hydrolases from the extreme thermophile Caldicellulosiruptor isolate Tok7B.1.";
Curr. Microbiol. 40:333-340(2000).
EMBL; AF078040; AAK06391.1; --
HSSP; Q06851; 1NBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                            Length 1770;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      / Match 40.7%; Score 191.5; DB 2; Length 261; Local Similarity 43.5%; Pred. No. 2.6e-13;
                                                                                31; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            32; Indels
PRINTS; PR01217; PRICHEXTENSN.
ProDom; PD001947; CBD_3; 3.
PROSITE; PS000591; GLYCOSYL, HVDROL_F10; 1.
SEQUENCE 1770 AA; 193641 MW; BBAF1937D4926C92 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                261
28759 MW; 4771744A26A6AE04 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                              Last sequence update)
Last annotation update)
                                                                                                                                                                                                                       09AG7;
01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-JUN-2001 (TrEMBLrel. 19, Last annotation update
GLYCOSYL HYDROLASE 6 (FRAGMENT).
GLAIGELLULOSITUPTOR SP. TOK7B.1.
Bacteria: Firmicutes; Bacillus/Clostridium group;
Thermoanaerobacter group; Caldicellulosiruptor.
                                                            DB 2;
                                                            ch 41.0%; Score 192.5; DB 2
1 Similarity 43.5%; Pred. No. 1.9e-12;
37; Conservative 16; Mismatches 31
                                                                                                                                                                                                                261 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 1426 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15; Mismatches
                                                                                                                                                        1163 IGASNVTFNFVKLSSGVSGADYYLE 1187
                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                      STRAIN=TOK7B.1;
MEDLINE=20171169; Pubmed=10706665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           177 IGASNVTFNFVKLSSGVSGADYYLE 201
                                                                                                                                           64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                            Interpro; IPR001956; CBD_3.
Pfam; PF00942; CBD_3; 1.
ProDom; PD001947; CBD_3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  12,
12,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (TrEMBLrel. 1 (TrEMBLrel. 1 (TrEMBLrel. 1
                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                           Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         261 AA;
                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                        NCBI_TaxID-80339;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1999 (
01-NOV-1999 (
01-DEC-2001 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                               09AQG7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              945X9Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        09x3P6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
Q9X3P6
                                                                                                                                                                                          RESULT
                                                                                                                                                                                                      Q9AQG7
  SO DR SO
                                                                                                                                           ò
                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
```

```
A Gibbs W.D., Receive R.A., Farrington G.K., Anderson P., Williams D.P.,
A Gibbs W.D., Receiver R.A., Farrington G.K., Anderson P., Williams D.P.,
Bergquist P.L.;
T "Multidomain and multifunctional glycosyl hydrolases from the extreme
the curr. Microbiol. 40:333-340(2000).
BERLA ROTROSTRIA AAD30364.1; -.
BERLA ROTROSTRIA GLYCO-hydro_10.
BERLA ROTROSTRIA GLYCO-HYDROL ROTROSTRIA ROTROSTRIA ACTINITALL IN UNKNOWN_1.
BEROSTRE, ROSO0099; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
BEROSTRE, ROSO0699; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
BEROSTRE, ROSO0699; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
BEROSTRE, ROSO0699; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
BEROSTRE, ROSO0699; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -!- CATALYTIC ACTIVITY: ENOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC LINKAGES IN CELLULOSE.
-!- SIMILARITY: THE N-TERMINAL PART BELONGS TO CELLULASE FAMILY E (FAMILY 9 OF GIYCOSYL HYDROLASES)
-!- SIMILARITY: THE C-TERMINAL PART BELONGS TO CELLULASE FAMILY L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae; Anaerocellum group; Anaerocellum. NCBI_TaxID=31899;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: THE N-TERMINAL DOMAIN OF CELA ENCODES FOR AN ENDOGLUCANASE ACTIVITY ON CARBOXYMETHYLCELLULOSE. THE C-TERMI DOMAIN PROBABLY ACT SYNERISTICALLY TO HYDROLYZE CRYSTALLINE CELLULOSE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 1426;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-MAY-1997 (TrEMBLrel. 03, Created)
01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
ENDOGLUCANASE A (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE A)
Bacteria; Firmicutes; Bacillus/Clostridium group; Thermoanaerobacter group; Caldicellulosiruptor. NCBL_TaxID-80339;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 40.7%; Score 191.5; DB 2; Best Local Similarity 43.5%; Pred. No. 1.9e-12; Matches 37; Conservative 15; Mismatches 32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 1711 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE=20171169; PubMed=10706665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (CELLULASE A) (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Anaerocellum thermophilum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
```

ï

1779 AA

S

```
4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN=Y106; O1 M., Wang P., Liu X., Qu Y.; Wang P., Liu S., Qu of a cellulase from Bacillus subtilis Y106.",
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-RT69B.1;
Moris D.D., Gibbs M.D., Ford M., Thomas J., Bergquist P.L.;
"Family 10 and 11 xylanase genes from Caldicellulosiruptor sp.
Rt69B.1.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AUDILITIES (DEC-1997) to the EMBL/GenBank/DDBJ databases.

REMBL; AF036924; AAB95326.1; -
HSSP; Q06851; 1NBC.

InterPro; IPR001000; G19C.

InterPro; IPR001305; CBD_6; 2.

InterPro; IPR00131; G19C.

InterPro; IPR00134; G19C.

INTERPRO; INTERPRO; INTERPRO; G19C.

INTERPRO; INTERPRO; INTERPRO; G19C.

INTERPRO; INTERPRO; INTERPRO; G19C.

INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO; INTERPRO
                                                                                                                                                                                                                                                                                                                                                   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
FAMILY 10 XYLANASE (EC 3.2.1.8).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 40.5%; Score 190.5; DB 2; Best Local Similarity 43.5%; Pred. No. 3.3e-12; Matches 37; Conservative 14; Mismatches 33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Caldicellulosiruptor sp. Rt69B.1.
Bacteria: Firmicutes; Bacillus/Clostridium group;
Thermoanearobacter group; Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Created)
                                                                                                                                                                                                                                                                                                                         Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1172 IGASNVTFNFVKLTSGVSGADYYLE 1196
                                                                                                                                                                                                                                                           PRT;
                                                                       903 IGASNVTFNFVKLSSGVSGADYYLE 927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                   64 MGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                   01-JUN-1998 (TrEMBLrel. 06, 01-JUN-1998 (TrEMBLrel. 06, 01-DEC-2001 (TrEMBLrel. 19,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19, ALKALI TOLERABLE CELLULASE.
                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bacillus subtilis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=1423;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q93TJ6
                                                                                                                                                                                                                                                     052374
                                                                                                                                                                                             RESULT 12
052374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                           g
                                                                                                                                                                                                                                                           TD DDT 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-TOK7B.1;
MEDLINE-2017169; Pubmed-10706665;
Glibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P.,
Bergquist P.L.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 VSGG-VKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNC 59
                                                              DR InterPro; IPR001956; CBD_3.

DR InterPro; IPR001956; CBD_3.

DR Ffam; PF001942; CBD_3; 3.

DR Pfam; PF001959; Glyco_hydro_9; 1.

PRINTS; PR001844; GLHYDRLASE48; 1.

PRODOM; PD011940; GLYCO_hydro_48; 1.

PROSITE; PS00592; GLYCOSYL_HYDROL_F9_1; 1.

PROSITE; PS00598; GLYCOSYL_HYDROL_F9_1; 1.

PROSITE; PS00698; GLYCOSYL_HYDROL_F9_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  40.6%; Score 191; DB 2; Length 1711;
43.8%; Pred. No. 2.7e-12;
Live 18; Mismatches 30; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3C72B6ED22F3C614 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
GLYCOSYL HYDROLASE 5 (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Caldicellulosiruptor sp. Tok78.1.
Bacteria; Firmicutes; Bacillus/Clostridium group; Thermoanaerobacter group; Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
40.5%; Score 190.5; DB 2;
Best Local Similarity 43.5%; Pred. No. 1.6e-12;
Matches 37; Conservative 15; Mismatches 32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            996 AA
(FAMILY 48 OF GLYCOSYL HYDROLASES).
EMBL; Z86105; CAB06786.1; -.
HSSP; P26221; 1TF4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               742 DWAQIGASNVTFKFVKLSSSVSGADYYLE 770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    60 DWAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Interpro; IPR002860; BNR.
Interpro; IPR001956; CBD_3.
Interpro; IPR002965; P_rich_extensn.
PFGan; PFC2012; BNR; 9.
PRNNTS; PR01217; PRICHEXTENSN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        108275 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ProDom; PD001947; CBD_3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        996 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
Matches 39; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=80339;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hydrolase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q9AQH0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9AQHO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
```

ä

Gaps

1;

499 AA.

Length 1779;

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                  7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ä
                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                  4; Gaps
                                                                                                                                                                                                                                                                      4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-79-23;
Jung K.H., Chun Y.C., Lee J.-C., Kim J.H., Yoon K.-H.;
"Cloning and expression of a Bacillus sp. 79-23 cellulase gene.";
Biotechnol. Lett. 18:1077-1082(1996).
                                                                                                                                                 40.4%; Score 190; DB 2; Length 499; illarity 40.7%; Pred. No. 8.1e-13; Conservative 21; Mismatches 26; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ch 40.4%; Score 190; DB 2; Length 508; I Similarity 40.7%; Pred. No. 8.3e-13; 35; Conservative 21; Mismatches 26; Indels
Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF355629; AAK39540.1; -.
SEQUENCE 499 AA; 55252 MW; 587A687584B7E5BE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-CH21;
Zvidzai C.J., Delgado O.D., Zvauya R., Mattiasson B.;
Zvidzai C.J., Delgado O.D., Zvauya R., Mattiasson B.;
"Bacillus subtilis CH21 endo-b-1,4-glucanase gene.";
Submitted (Jul.-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; Av044252; AAK94871.1;
Hydrolase; Glycosidase.
SEQUENCE 508 AA; 56553 MW; A395EEICGF3340F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          093LD0;
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
ENDO-1, 4 - BETA-GLUCANASE (EC 3.2.1.4).
Bacillus subtilis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CELS.
Bacillus sp. 79-23.
Bacteria: Firmicutes: Bacillus/Clostridium group:
Bacillus/Staphylococcus group: Bacillus.
NCBI_TaxID=72363:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
NCBI_TaxID=1423;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        499 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Created)
                                                                                                                                                                                                                                                                                                                                                                                                                             :||||: | ||||: | | |||||: | ||||||: | 410 QIGCGNVTHKFVTLHKPKQGADTYLE 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       419 QIGCGNVTHKFVTLHKPKQGADTYLE 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                           63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ENDO-B-1,4-GLUCANASE (EC 3.2.1.4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-JUN-1998 (TrEMBLrel. 06, 01-JUN-1998 (TrEMBLrel. 06, 01-DEC-2001 (TrEMBLrel. 19,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
Matches 35; Conserva
                                                                                                                                                 Query Match
Best Local Similarity
Matches 35; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 14
093LD0
C 093LD0
DT 01-DEC
DT 01-DEC
DT 01-DEC
DE ENDO-1
OS BACILLI

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     052731
                                                                                                                                                                                                                                                                                                                                                                                           δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SPR
                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

THIS PAGE BLANK (USPTO)

C. thermocellum ce Amino acid sequenc Amino acid sequenc 60 kD endoglucanas Endoglucanase enco Cellulose binding

Scaffoldin protein C. thermocellum ce

C. cellulovorans C. Colstridium cellul C17E2 Osph Constru Chimeric S peptide Clostridium cellul Collulovorans C. cellulovorans C. cellulovorans C. cellulovorans C. cellulovorans C. cellulovorans C. human brancherin colluman G-protein colluman Housel Lagnent Housel Clost Fragment Housel Clost Fragment Housel Colluman Clost Fragment Housel Clost Fragment Housel Clost Fragment Housel Colluman G-protein Colluman G-protein Colluman G-protein Colluman G-protein Colluman Clost Fragment Housel Clost Fragment Fra

Single-chain anti-Human p53 regulato Protein sequence o Murine monoclonal

Orosophila melanog

from the

231.5 195.5 194.5 191.5 186 172 165.5 165 165

10 7 10 10 11 11

Score

Result No.

Sequence of the si Anti-erbB2 scFv.

Nasturtium xyloglu

Perfect score:

Run on:

Scoring table: Sequence:

Searched:

Database

```
Cellulase AE-1 for e.g. mfr. of pharmaceuticals and foodstuffs of opt. pH when carboxy:methyl:cellulose is used as substrate.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Detergents; pharmaceuticals; deinking; carboxymethylcellulose.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The sequence was deduced from the gene which was sequenced plasmid, pAEC 1, prepd. by ligating chromosomal DNA contg.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ALIGNMENTS
                                                                                                                               AAE05745
AAB81128
AAE05749
AAE05747
                                                                                                                                                                                                               AAE05746
AAY39952
                                                                                                                                                                                                                                                 AAW90081
AAW90080
AAY40440
                                                                                                                                                                                                                                                                                                AAU04567
AAU04581
AAU04570
ABG24680
AAU17759
                                                                                                                                                                                                                                                                                                                                                                                                  AAW56445
AAW56441
AAR95287
                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAR45442
AAR94020
AAW15238
AAW43108
AAG63962
AAG63963
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAW86130
                                                                                                                                                                                                                                                                                                                                                                                    AAW56450
                                                                                                                 AAW90077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAR15625 standard; Protein; 782 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 2; Fig 3; 8pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 90JP-0045465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  90JP-0045465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            17-MAR-1992 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               212
(OJIP ) OJI PAPER KK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Aeromonas strain no.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1991-373412/51.
N-PSDB; AAQ15178.
Cellulase AE-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JP03251174-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28-FEB-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-FEB-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                08-NOV-1991.
AAR15625;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR15625
 RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Cellulase AE-1. A
Truncated cellulas
Truncated cellulas
Truncated cellulas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1980.DAT:*
3: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1981.DAT:*
4: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1981.DAT:*
5: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1981.DAT:*
6: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1981.DAT:*
7: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1981.DAT:*
8: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1980.DAT:*
9: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1980.DAT:*
10: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1980.DAT:*
11: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1980.DAT:*
12: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1980.DAT:*
13: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1990.DAT:*
14: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1991.DAT:*
15: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1991.DAT:*
16: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1991.DAT:*
17: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1991.DAT:*
18: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1991.DAT:*
19: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1999.DAT:*
19: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1999.DAT:*
19: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1999.DAT:*
19: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1999.DAT:*
20: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1999.DAT:*
22: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1999.DAT:*
22: /SIDS1/gcgdata/hold-geneseqy/geneseqp-embl/AA1999.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NK-1 cellulase. B
Novel endoglucanas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     endoglucanas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Corrected Bacillus
Cellulose binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A mannanase-linker
Pectate lyase-link
Pectate lyase CBD
                                                                                                            2, 2002, 09:15:51; Search time 101.03 Seconds (without alignments) 96.748 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
                                                                                                                                                                                                               1 VSGGVKVQYKNNDSAPGDNQ.....IRASFGSVNPATPTADTYLQ
             GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                               Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                  747574 seqs, 111073796 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                 Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR15625
AAY13494
AAY13493
AAY13493
AAR13492
AAR13227
AAR13227
AAR13227
AAR1500
AAY54123
AAY54123
AAY54123
                                                                                                                                                                                                                                                  BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A_Geneseq_032802:*
                                                                                                                                                                                                                                                                                                                                                                Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-917-376-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1751
1426
499
700
551
167
476
493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               335.22
35.22
35.22
35.22
35.22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221: ... 221
```

a

```
1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The invention relates to a recombinant cellulase active protein free of proteinases of native thermophilic and alkalinophilic origin, comprising the truncated sequences cel B5, Cel B1, Cel E1, Cel E1/2, Cel I1/2), Cel E1/3, Cel E 10/2, Cel E1/2, 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Cellulase; proteinase; truncated; Cel B5; Cel B4/5; Cel E1, Cel E1/2; Cel 1/2/3; Cel 6; Cel E3/B5; laundry detergent; stain-removing; cotton-containing fabric; stonewashing.
                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                        2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
gene (obtd. from Aeromonas) into pUC18. The protein has amol. wt. of 81.000 (SDS-PAGE) and an optimum pH near to 5 when carboxymethylcellulose is the substrate. The Yerminal sequence: GIHADT has been confirmed by Edman degradation. The gene can be used to produce recombinant enzyme which is used for the effective utilisation of biomass resources and the mfr. of pharmaceuticals and foodstuffs, and also for the detergent and deinking of waste paper.
                                                                                                                                                                                                                                                                                                                                                      1;
                                                                                                                                                                                                                                                                                                 Length 782;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New truncated cellulase proteins, useful in detergents and for producing 'stonewashed' denim
                                                                                                                                                                                                                                                                                                                                                   28; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Farrington GK;
                                                                                                                                                                                                                                                                                                 DB 12;
                                                                                                                                                                                                                                                                                              49.4%; Score 231.5; DB 48.3%; Pred. No. 5e-18; tive 16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Daniels RM,
ms DP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 7; Page 42-43; 65pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY13494 standard; Protein; 616 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (CLRN ) CLARIANT FINANCE BVI LTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             t PL, Dan
Williams
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Truncated cellulase Cel E3/B5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97US-0932571.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98EP-0810919.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                      42; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Anderson P, Bergquist
Gibbs MD, Morgan H, W
                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1999-315403/27.
                                                                                                                                                                                                                   782 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-PSDB; AAX55660.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 15-SEP-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-SEP-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        30-JUL-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EP921188-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09-JUN-1999
                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY13494;
                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                          Best Local
                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY13494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
  888888888
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

```
ï
remove staining, backstaining or graying, for use on cellulosic materials including cotton-containing fabrics. They are especially useful for preventing redeposition of colorant during stonewashing, and for processing of textiles where cellulose breakdown is required. The new truncated enzymes show reduced redeposition of dye compared to using non-truncated cellulase compositions.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The invention relates to a recombinant cellulase active protein free of proteinases of native thermophilic and alkalinophilic origin, comprising the truncated sequences Cel B5, Cel B4/5, Cel E1/2, Cel I1/2/3, Cel Col Gor Cel E3/B5, or a stability region from one of the defined fulllength sequences, or functional equivalents. Cel B5 extends from amino acid A1011 to P1424 or K1425 or N1426, and Cel B4/5 extends from amino acid K635 to N1426 in the sequence shown in AAY13492; Cel E1 extends from amino acid Y39 to G812, Cel E1/2 extends from 39 to G635, Cel E1/2/3 extends from 139 to G812, Cel E6 extends from amino acid V1233 to K1751 and the stability region extends from amino acid W1233 to the sequence shown in AAY13493; Cel E3/B5 is shown in AAY13494. The new enzymes are useful in laundry detergent compositions to prevent or remove staining, backstaining or graying, for use on cellulosic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Cellulase; proteinase; truncated; Cel B5; Cel B4/5; Cel E1; Cel E1/2; Cel 1/2/3; Cel 6; Cel E3/B5; laundry detergent; stain-removing; cotton-containing fabric; stonewashing.
                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                               1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                 ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 for
                                                                                                                                                                              20; Length 616;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               proteins, useful in detergents and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Daniels RM, Farrington GK;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Truncated cellulases comprising amino acid sequence.
                                                                                                                                                                            ; Score 195.5; DB 20; Pred. No. 5.4e-14; 15; Mismatches 34
                                                                                                                                                                                                                                                                                                                    61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                   waqigasnvtfnfvklssgvsgadyyle 87
                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY13493 standard; Protein; 1751 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 7; Page 37-41; 65pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Williams DP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (CLRN ) CLARIANT FINANCE BVI LTD.
                                                                                                                                                                            41.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   denim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        97US-0932571.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98EP-0810919.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bergquist PL,
                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 New truncated cellulase producing 'stonewashed'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Morgan H,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1999-315403/27
                                                                                                                                                                                            Similarity
                                                                                                                         616 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-PSDB; AAX55662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15-SEP-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19-SEP-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Anderson P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EP921188-A2.
                                                                                                                                                                            Query Match
Best Local Simi
Matches 38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            30-JUL-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gibbs MD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY13493;
                                                                                                                           Seguence
                                                                                                                                                                                                                                                                                                                                                     09
                                                                                                                                                                                                                                                                                                                                                                                                                    AAY13493
     8888888888
                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                     q
```

 $\sim$ 

888888888

ò g ò 셤

```
Bacillus N4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Simmatches 35;
                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAR42122;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                         Query Match
                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
AAR42122
                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QΩ
      88888
                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δλ
                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The invention relates to a recombinant cellulase active protein free of proteinases of native thermophilic and alkalinophilic origin, comprising the truncated sequences Cel B5, Cel B4/5, Cel B1/2, Cel 11/2/3, Cel 6 or Cel 13/85, or a stability region from one of the defined full-length sequences, or functional equivalents. Cel B5 extends from amino acid A1011 to P1424 or K1425 or N1426, and Cel B4/5 extends from amino acid K635 to N1426 in the sequence shown in AAY13422, Cel E1 extends from amino acid K139 to D481, Cel E6 extends from 39 to G635, Cel E1/2, axtends from 439 to G812, Cel E6 extends from amino acid W1233 to K1751 and the stability region extends from amino acid R482 to G635, in the sequence shown in AAY13493, Cel E6 extends from amino acid W1233 to K1751 and the stability region extends from amino acid E482 to G635, in the sequence shown in AAY13493, Cel E3/85 is shown in AAY13494. The new enzymes are useful in laundry detergent compositions to prevent or remove staining, backstaining or graying, for use on cellulosic materials including cotton-containing fabrics. They are especially useful
                                                                                                                                                                                                                                  1;
materials including cotton-containing fabrics. They are especially useful for preventing redeposition of colorant during stonewashing, and for processing of textiles where cellulose breakdown is required. The new truncated enzymes show reduced redeposition of dye compared to using non-truncated cellulase compositions.
                                                                                                                                                                                                                                                                                              Cel B4/5; Cel E1; Cel E1/2;
                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                         GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                  ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Cellulase; proteinase; truncated; Cel B5; Cel B4/5; Cel E1; Cel Cel 1/2/3; Cel 6; Cel E3/B5; laundry detergent; stain-removing; cotton-containing fabric; stonewashing.
                                                                                                                                                                                          Length 1751;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               New truncated cellulase proteins, useful in detergents and for producing 'stonewashed' denim
                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Farrington GK;
                                                                                                                                                                                        DB 20;
                                                                                                                                                                                                                                32;
                                                                                                                                                                                     Query Match 41.5%; Score 194.5; DB 2
Best Local Similarity 44.7%; Pred. No. 2.6e-13;
Matches 38; Conservative 14; Mismatches 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Truncated cellulases Cel B4/5 and Cel B5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Anderson P, Bergquist PL, Daniels
Gibbs MD, Morgan H, Williams DP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 7; Page 34-37; 65pp; English.
                                                                                                                                                                                                                                                                                                                                                                                            737 igasnvtfnfvklssgvsgadyyle 761
                                                                                                                                                                                                                                                                                                                                                        64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY13492 standard; protein; 1426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (CLRN ) CLARIANT FINANCE BVI LTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98EP-0810919.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   97US-0932571.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1999-315403/27
                                                                                                                            1751 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; AAX55661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             15-SEP-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-SEP-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    30-JUL-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EP921188-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY13492;
                                                                                                                                                                                                                                                                           4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY13492
```

```
ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 for preventing redeposition of colorant during stonewashing, and for processing of textiles where cellulose breakdown is required. The new truncated enzymes show reduced redeposition of dye compared to using non-truncated cellulase compositions.
                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                      4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequences (AAQ49818-19) consist of two synthetic fragments which are
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    used to induce a mutation within the cellulase NK-1 gene. The gene is shown in sequence (AAQ49820). The modified NK-1 gene shows a change in pH dependence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Changing the pH-dependence of cellulase enzymatic activity - by changing base sequence of cellulose-producing gene of Bacillus microbe to base sequence coding asparagine and serine residues at
                                                                                                                                                                                             Length 1426;
                                                                                                                                                                                                                                         ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 499;
                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                        DB 20;
                                                                                                                                                                                                                                         32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              39.7%; Score 186; DB 14; 40.7%; Pred. No. 5.2e-13;
                                                                                                                                                                                        40.8%; Score 191.5; DB 20
43.5%; Pred. No. 4.4e-13;
live 15; Mismatches 32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ā
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Cellulase; pH dependence; mutation.
                                                                                                                                                                                                                                                                                                                                                                                                                 :| |: :| :: || |: 472 igasnvtfnfvklssgvsgadyyle 496
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                   64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Disclosure; Fig 1; 9pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR42122 standard; Protein; 499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   92JP-0075883
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    92JP-0075883
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        specified aminoacid sites
                                                                                                                                                                                                                                         37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 1993-330585/42.
                                                                                                                   1426 AA;
                                                                                                                                                                                                                 Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAQ49820.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (BEPP/) BEPPU T.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NK-1 cellulase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JP05236969-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-FEB-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   28-FEB-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              27-APR-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     17-SEP-1993.
```

AAR13227;

AAR13227

RESULT

a

Peptide Protein

Key

```
2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSYDLSTVTVRYWFTRDGGSSTLVYNCDW,61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The present sequence is the corrected version of the incorrect Bacillus lautus (NCIMB 40250) endoglucanase Endo 3A described in W09110732. Endo 3 can be used in novel method of forming localised colour density variation on the surface of a dyed cellulosic fabric. The method comprises agitating the fabric in an aqueous medium (PH 6.5 to 9.0) containing a family 5 cellulose, e.g. the present sequence, which can hydrolyse p-nitrophenyl-beta abrading apent or cellulose having abrading activity. Each process, is useful to provide a stone washed look to blue jeans
                                                                                                                                                                                                                                                                                    Endoglucanase; Endo 3A; formation; localised; variation; colour density; surface; dye; fabric; family 5; cellulose; hydrolysation; p-nitrophenyl-beta-1,4-cellobioside; stone wash; blue jeans; back staining.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Localised variation of colour density in the surface of a dyed cellulosic fabric - uses cellulase compsn. able to hydrolyse
                                                                                                                                                                                                                         Corrected Bacillus lautus (NCIMB 40250) endoglucanase Endo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   35.3%; Score 165.5; DB 1
37.9%; Pred. No. 1.4e-10;
Live 22; Mismatches 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Toft AH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Disclosure; Pages 15-17; 23pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       p-nitrophenyl -beta-1,4-cellobioside
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | :| |: |:| | ||||::
461 aqigasnvsaaf--anftgsntdtyve 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AA.
                                      AAW18790 standard; protein; 551 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Fich M, Onishi M, Schulein M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAR95080 standard; peptide; 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  96WO-DK00364.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            95DK-0000993
                                                                                                                                                           18-NOV-1997 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        30-OCT-1996 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 37.94
Matches 33; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (NOVO ) NOVO-NORDISK AS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 without back staining
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1997-192888/17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           551 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                         Bacillus lautus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO9709410-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  03-SEP-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            08-SEP-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13-MAR-1997.
                                                                                              AAW18790;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR95080;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR95080
         AAW18790
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ω×
                                                                                                    X D X D X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X D X 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 XYXX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 VSGGVKVQYKNND-SAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNC 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The enzyme is produced by a strain of Bacillus spp. NCIMB 40250 and exhibits an endoglucanase activity of at least 10 (pref. at least 25) carboxymethyl cellulose (CMC) endoase units per mg total protein under alkaline conditions. It is especially useful as a cellulolytic agent and has been found to be more stable during washing (60 mins. at 40 deg.) in the presence of conventional detergents than a commercial cellulase preparation. It may also show increased storage stability in liq. detergents contg. proteases. The sequence was deduced from the DNA (AAQ13001). See also AAR13228 and AAR13229.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Enzyme exhibiting cellulase activity from Bacillus sp. - is an endo-glucanase, esp. useful for harshness redn. of cotton-contg.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             36.7%; Score 172; DB 12; Length 700; 44.9%; Pred. No. 3.3e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       27; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1..31
/label= signal sequence
32..700
/label= mature endoglucanase
31..32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ||| :| ||: |||: | | :||| :
| dwaklggsniqisfgnhnga--dsdtyae 634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             60 DWAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hansen C;
410 gigcgnlthkfvtlhkpkggadtyle 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                                                                        AAR13227 standard; Protein; 700 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Claim 1; Page 80; 96pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                        Cellulase activity; detergent.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  91WO-DK00013.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         90DK-0000164.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Jorgensen PL, Schulein M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          40; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacillus spp. NCIMB 40250.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (NOVO ) NOVO NORDISK A/S.
                                                                                                                                                                                                                                                                                                                                              Novel endoglucanase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1991-238020/32.
N-PSDB; AAQ13001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               700 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Cleavage-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18-JAN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19-JAN-1990;
                                                                                                                                                                                                                                                                                14-OCT-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WO9110732-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25-JUL-1991
```

fabrics

ä

Gaps

3;

Indels

29;

Sequence

Matches

ŏ g ò RESULT

DB 18; Length 551;

```
5;
                                                                                                                                                                                                                                                                                                                                                                                                to a hapten molety via one or more cys or Lys residues and complexes comprising the biotinylated CBD and a biotin-binding molecule selected from modified or unmodified avidin or streptavidin or an anti-biotin antibody, can be used in affinity chromatography, cell separation, cell, protein and enzyme immobilisation, selective removal of biological materials, multienzyme reactors, signal immunoassays and drug (e.g. pesticide) delivery. The CBD molecule can be modified with the hapten without affecting its high affinity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ||| :||:: |:: : | | | :: |||||::||| || 5 vsgnlkvefynsnpsdttnsinpqfkvtntgssaidlskltlryyytvdgqkdqtfw-cd 63
                    Cellulose binding domain; CBD; hapten; moiety; biotin; avidin; streptavidin; affinity chromatography; cell separation, cell immobilisation; protein immobilisation; enzyme immobilisation; multienzyme reactors; signal immunoassays; drug delivery; pesticide;
                                                                                                                                                                                                                                                                                                                                                                                         A modified cellulose binding domain (CBD) or fraction of it, linked
                                                                                                                                                                                                                                                                                                                        Modified cellulose-binding domain mols. - having attached hapten, partic. biotin, to provide an affinity system for, e.g. sepns., assays, reactors, delivery etc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 167;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A mannanase-linker-cellulose binding domain fusion protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                 Lamed R, Morag E, Shoham Y, Wilchek M;
                                                                                                                            /note= "Unidentified amino acid."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         35.2%; Score 165; DB 17;
36.1%; Pred. No. 3.6e-11;
ive 23; Mismatches 29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ||: :| ||||||| 64 haaiigsngsyngitsnvkgtfvkmssstnnadtyle 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 WAAM -----GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                     (TECR ) TECHNION RES & DEV FOUND LTD.
(UYRA-) UNIV RAMOT APPLIED RES & IND DEV LTD.
(YEDA ) YEDA RES & DEV CO LTD.
                                                                                                                                                                                                                                                                                                                                                                     Page 34-35; 53pp; English.
                                                                                                          Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY54123 standard; Protein; 476
                                                                                                                                                                                                                 94IL-0111415.
                                                                                                                                                                                              95WO-US13813.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match
Best Local Similarity 36.1%
Matches 35; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
Cellulose binding domain.
                                                                                    Clostridium thermocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       or cellulose, or chitin
                                                                                                                                                                                                                                                                                                     WPI; 1996-239453/24.
                                                                 cellulose; chitin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            167 AA;
                                                                                                           Key
Misc-difference
                                                                                                                                                   WO9613524-A1.
                                                                                                                                                                                                                  27-OCT-1994;
                                                                                                                                                                                            26-OCT-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  27-MAR-2000
                                                                                                                                                                        09-MAY-1996
                                                                                                                                                                                                                                                                                 Bayer EA,
                                                                                                                                                                                                                                                                                                                                                                     Claim 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY54123;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY54123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  XEXEXEX
```

```
Mannanase; mannan endo-1,4-beta-mannosidase; beta-mannanase; endo-1,4-mannanase; Bacillus sp. 1633; galactomannan; ando-1,4-beta-p-mannosidic linkage; mannan; galactomannan; glucomannan; galactoglucomannan; cellulosic fibre; synthetic fibre; yarn; fabric; printing paste; plant material degradation; recycled waste paper; paper making pulp; guar; locust bean gum; thickener; viscosity; mannan containing food; coffee extract; cleaning composition; machine washing; hard-surface cleaner; dishwashing; oral; dental; contact lens; body-care composition; fabric softener; oil well drilling; subterranean formation fracture; cellulose binding domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    endo-1,4-beta-mannosidase, beta-mannanase, or endo-1,4-mannanase)
hydrolyses galactomannans. Specifically, mannanases hydrolyse
hydrolyses galactomannans. Specifically, mannanases hydrolyse
l,4-beta-D-mannosidic linkages in mannans, galactomannans, glucomannans,
and galactoglucomannans. The mannanase protein, or preparations
containing it, are used to improve properties of cellulosic or
synthetic fibres, yarn or (non)woven fabrics (removal of mannan-based
sizes or printing pastes). They are also used to degrade or modify
plant material containing guar or locust bean gums (thickeners), or to
reduce viscosity of mannan-containing foods or feeds). The mannanases
are also used to process coffee extracts (to inhibit gel formation);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           in cleaning compositions (for machine washing of fabrics, as hard-surface cleaners, for hand or machine dishwashing, also in oral, dental, contact lens or body-care compositions) where they remove mannan-containing soils and prevent binding of some soils to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The present sequence represents a mannanase-linker-cellulose binding domain fusion protein. Mannanase (also known as mannan
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New mannanases for treatment of textiles, plant material and coffee
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bjornvad ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        cellulosics; and in fabric softeners. They can also be used in oil well drilling to fracture subterranean formations.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 476;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Schnorr K, Andersen LN,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 165; DB 21;
Pred. No. 1.3e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Example 4; Page 211-212; 242pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      extract, and in cleaning compositions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       35.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Schuelein M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99WO-DK00314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0105970
98US-0106054
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98DK-0001725
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99DK-0000306
99DK-0000307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99DK-0000308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0123641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98US-0111256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98DK-0001340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98DK-0001341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99DK-0000309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99US-0123543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0123623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0123642
                                                                                                                                                                                                                                                                                                                                                                                          Bacillus sp.
Clostridium thermocellum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (NOVO ) NOVO-NORDISK AS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2000-105891/09.
N-PSDB; AAZ45336.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           476 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Kauppinen MS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO9964619-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          10-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-OCT-1998;
28-OCT-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                23-DEC-1998;
05-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  05-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            05-MAR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           09-MAR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10-JUN-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         20-OCT-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20-0CT-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0-MAR-1999
                                                                                                                                                                                                                                                                                                                                                             Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
```

us-09-917-376-5.rag

A.

q ð g

å

```
AAY43218 standard; Protein; 493
                                                                                                                                                 13-JAN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Glad SOS,
                                                                                                                                                                                                                                                                                                                                Synthetic.
                                                                                                                    AAY43218;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                          RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Loca
Matches
               g
                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ?
 ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Plasmid pMB914 was constructed using pectate lyase gene from Bacillus licheniformis, ATCC 14580 and cellulose binding domain (CBD) gene from Clostridium thermocellum Ys fused with a linker molecule.

Bacillus subtilis was transformed with plasmid pMB914 for expression of the fusion protein. Pectate lyase can be used in detergent compositions, for cleaning hard surfaces, for machine treatment of fabrics, for improving the properties of cellulosis fibres, yarn, woven or non-woven fabric, for the degradation of plant material e.g. recycled waste paper, mechanical paper making pulps or fibres subjected to retting process, for preparing animal feed and for processing wine or juice.
                                                                                                                                                                                                                                                                                                              Pectate lyase-linker-CBD fusion protein; ATCC 14580; plasmid pMB914; cellulose binding domain; CBD; linker; transformed; detergent; cellulosic fibre; yarn; degradation; recycled waste paper; animal feed; paper-making pulp; retting process; processing; wine; juice.
                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The present sequence is a Pectate lyase-linker-CBD fusion protein.
                             VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Lange NEK, Schnorr K, Schuelein M;
 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 493;
29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            35.2%; Score 165; DB 20; 36.1%; Pred. No. 1.4e-10; iive 23; Mismatches 29;
                                                                                                         61 WAAM ------GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----GCGNIRASFGSVNPATPTADTYLQ 88
23; Mismatches
                                                                                                                                                                                                                                                                                   Pectate lyase-linker-CBD fusion protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Example 4; Page 85-86; 93pp; English.
                                                                                                                                                                                                                                                                                                                                                                                        Chimeric - Bacillus licheniformis.
Chimeric - Clostridium thermocellum.
                                                                                                                                                                                           AAY28850 standard; Protein; 493 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New isolated pectate lyase enzymes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98US-0073684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                98WO-DK00514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            97DK-0001344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bjornvad ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (NOVO ) NOVO-NORDISK AS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1999-610578/52.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity
nes 35; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   493 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; AAX90978
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Andersen LN,
                                                                                                                                                                                                                                                                                                                                                                                                                                     W09927083-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          06-MAY-1998;
                                                                                                                                                                                                                                                      17-JAN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24-NOV-1997;
35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WAAM--
                                                                                                                                                                                                                        AAY28850;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                             AAY28850
```

a ò

õ

```
ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This sequence is a fusion protein comprising a Bacillus species pectate lyase of the invention, fused via a linker to the Clostridium thermocellum CipB protein cellulose binding domain (CBD). The pectate lyase enzymes are obtained from Bacillus agaradhaerens (BA), Bacillus licheniformis (BL), Bacillus halodurans (BH) and other Bacillus species. The pectate lyase enzymes can be used for degrading pectin, pectate and polygalacturonides. They can be used in detergent compositions, for cleaning hard surfaces, for machine treatment of fabrics, for improving the properties of cellulosic fibres, yarm, woven or non-woven fabric, for the degradation of plant material (e.g. recycled waste paper, mechanical paper-making pulps or fibres subjected to a retting process, for
                                                         Pectate lyase; polysaccharide lyase; enzyme; pectin degradation; polygalacturonide; detergent composition; hard surface treatment; cellulosic fibre; plant material degradation; recycled waste paper; mechanical paper-making pulp; wine processing; cellulose binding domain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              preparing animal feed or for processing wine or juice. DNA encoding the enzymes can also be used for the production of transgenic plants.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Moller
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 493;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bjornvad ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Lange NEK, Bjornvad M. Schnorr K, Kongsbaek L;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            35.2%; Score 165; DB 20;
36.1%; Pred. No. 1.4e-10;
iive 23; Mismatches 29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
Pectate lyase CBD fusion protein sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Claim 26; Page 106-108; 113pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New isolated pectate lyase enzymes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         97DK-0001344.
98US-0073684.
98US-0184217.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98WO-DK00515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              97DK-0001343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Schuelein M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Kauppinen MS,
                                                                                                                                                                                                                                                                  Clostridium thermocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (NOVO ) NOVO-NORDISK AS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1999-610579/52.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity
nes 35; Conserv.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAZ31562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Andersen LN,
                                                                                                                                                                                                                                                                                                                                                                                                                                                             03-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24-NOV-1997
06-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             02-NOV-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

531 AA;

Seguence

g

용

```
This protein sequence represents a scaffoldin based on the CipA protein of Clostridium thermocellum. The scaffoldin protein is used in a novel composition that comprises at least 2 enzymes non-covalently bound to a peptide backbone (i.e. present sequence). The scaffoldin comprises a number of internal repeating units and at least one cellulose binding domain (CBD). The CBD may be altered to modify its affinity for cellulose, which may be desirable where cellulose binding would be disadvantageous. The enzyme is bound to the scaffoldin by a dockerin region of the enzyme, which binds to a repeating element of the scaffolding. The dockerin is preferably (CELS (AAW15237) or CelD (AAW15236). The composition can be used in reducing allergenicity, producing synergistic effects, and facilitating selective modification of substrate. By taking advantage of the cellulose binding domain of the complex, the complex could be immobilised for use in chromatographic scaffolding domain, it is possible to recover enzymes, or to quantify the amount of an enzyme in a solution. The composition could also be used in a multi-enzyme delivery system which could be used in the food industry, in food processing, animal feed, textiles, bioconversion, pulp and paper production, plant protection and pest control, as a wood preservative, topical lotions, and biomass conversions.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Composition with enzymes non-covalently bound to a peptide backbone - used as a multi-enzyme delivery system, e.g. in food processing,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Dockerin; CelD; CelS; CipA; scaffoldin; cellulose binding domain; chromatographic separation; soluble substrate modification; CBD; multi-enzyme delivery system; animal feed; paper production; plant protection; pest control.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Xia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /label= internal_repeat_element_2
239..531
/label= cellulose_binding
/note= "only 60% of the CBD is present"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ward M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1..153
/label= internal_repeat_element_1
                                                                                                                                                                                                                                                                                                                                                                                                                 Scaffoldin protein from Clostridium thermocellum.
Liu C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Location/Qualifiers
                                                                                                                                                                                                      AAW15238 standard; protein; 531 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Clarkson KA, Fowler T,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 10; Fig 6; 29pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   95US-0559968.
95US-0005701.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      96WO-US16485
                                                                                                                                                                                                                                                                                                                                        28-JAN-1998 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           textiles and pest control
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Clostridium thermocellum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  154..306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GEMV ) GENENCOR INT INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 1997-245106/22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   17-NOV-1995;
17-OCT-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO9714789-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16-OCT-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 24-APR-1997
                                                                                                                                                                                                                                                                           AAW15238;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bott RR,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Key
Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Domain
                                                                                                                                         12
                                                                                                                                                                      AAW15238
                                                                                                                                         RESULT
```

```
wultimeric protein, especially enzymatic, complexes are held together
by protein-protein interactions between domains designated dockerins
and cohesins, which are found on the catalytic and scaffold subunits
respectively. An example of such a complex is the cellulose degrading
protein complex from clostridium thermocellum, known as the cellulosome.
This complex comprises around 15 proteins including endoglucanases,
cellobiohydratases, hemicallulases, eg. xylanases or lichinases, which
interact with a central "scaffold" protein designated the cellulosome
integrating protein (Cip; see AAW43108). The catalytic subunits
integrating protein (Cip; see AAW43108). The catalytic subunits
integrating rotein (Cip; see AAW43108).
The invention relates to the isolation of proteins binding to a novel
dockerin type domain found in the C-terminal portion of Cip. The new
domain is designated a type II dockerin domain (as compared to the type
II dockerin domain has some sequence similarity to the type I dockerins
but is unable to bind type I cohesin domains.
                                      3
                                                                                                                                                                                                                                                                                                                                                                                                                                            Multimer; enzyme; complex; protein-protein interaction; dockerin domain; cohesin domain; catalytic subunit; scaffold subunit; SdbA; synergistic; cellulosome integrating protein; scaffoldin dockerin binding protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The sequence presented here is the cellulosome integrating protein (CIP-A) to which the proteins SdbA, OlpB and ORF2P (AMW4105-W43107) can bind at the C-terminal dockerin type II domain (location not given in specification). The novel type II dockerin and cohesin domains can be
                                                                                                 Gaps
                                                                          1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Cellulase proteins with cohesin or dockerin type II domains – useful for potentiating the activity of multiprotein enzyme complexes
                                      10;
  Length 531;
                                        Indels
Score 165; DB 18;
Pred. No. 1.5e-10;
3; Mismatches 29;

    C. thermocellum cellulosome integrating protein.

                                                                                                                                                                               ||: | |||||: 371 haaiigsngsyngitsnvkgtfykmssstnnadtyle 407
                                                                                                                                                       61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
                                    23; Mismatches
                                                                                                                                                                                                                                                                                         AAW43108 standard; Protein; 1853 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Page 47; 60pp; French.
  35.2%;
36.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            96FR-0005854
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 96FR-0005854
                                                                                                                                                                                                                                                                                                                                                                   16-OCT-1998 (first entry)
                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Clostridium thermocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ü
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (INSP ) INST PASTEUR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Beguin P, Leibovitz
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1998-011569/02.
  Query Match
Best Local Similarity
Matches 35; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAT86625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FR2748479-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10-MAY-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10-MAY-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      14-NOV-1997.
                                                                                                                                                                                                                                                                                                                               AAW43108;
                                                                              ò
                                                                                                                 g
                                                                                                                                                       δλ
                                                                                                                                                                                          a
```

us-09-917-376-5.rag

Thu Jul 18 16:01:12 2002

ï

1; Gaps

Indels

Pred. No. 5.9e-09; 1; Mismatches 36;

21; 31.8%;

Best Local Similarity 31.8% Matches 27; Conservative

à g δλ

4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63

```
5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   specification describes a xyloglucanase enzyme belonging to family 44 of at least 10% at per second exhibiting a relative xyloglucanase activity of at least 10% at pH 5-8. The enzyme exhibits high performance in detergent compositions and prevents binding of certain soils to the xyloglucan left in the cellulosic material. It is stable at pH 5-10 at room temperature and has a half life of more than 50 days when incubated in a full formulated liquid detergent at 30 plus degrees celsius. The enzyme is used in detergent compositions, textile industry for improving the properties of cellulosic fibers, yarn, and woven or non-woven fabrics, preferably in textile scouring process, and in cellulose fiber processing industry for ratting of fibers e.g. hemp, jute, flax, and
                                                                                                                                                                                           The present sequence represents a xyloglucanase of the invention. The
                                                                                                                                        Gaps
used in complexes, especially enzyme complexes, to potentiate their catalytic actions in a synergistic manner.
                                                                                                                                                                        VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    New xyloglucanase enzyme belonging to glycosyl hydrolases family, useful for detergent compositions, and textile or cellulose fiber
                                                                                                                                      10;
                                                                                                   Length 1853;
                                                                                                   ; Score 165; DB 19; Length 18; Pred. No. 7e-10; 23; Mismatches 29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          xyloglucanase; family 44; glycosyl hydrolase; detergent; cellulosic fiber; textile scouring.
                                                                                                                                                                                                                                                             ||:::| :::| |||||:
424 haailgsngsyngltsnvkgtfvkmssstnnadtyle 460
                                                                                                                                                                                                                                        61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Amino acid sequence of xyloglucanase enzyme.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Schuelein M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Example 3; Page 81-85; 97pp; English.
                                                                                                                                                                                                                                                                                                                                                           AAG63962 standard; Protein; 1352 AA
                                                                                                    35.2%;
36.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21-FEB-2001; 2001WO-DK00116.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 24-FEB-2000; 2000DK-0000291
                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                               Query Match
Best Local Similarity 36.19
Matches 35; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Schnorr K, Jorgensen PL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Paenibacillus polymyxa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     processing industries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (NOVO ) NOVOZYMES AS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2001-522819/57.
N-PSDB; AAH75059.
                                                 1853 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1352 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200162903-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                            29-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 30-AUG-2001
                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                            AAG63962;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                              AAG63962
 SSXS
                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                        δ
```

Length 1352;

33.2%; Score 155.5; DB 22;

Query Match

```
ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 The present sequence represents a xyloglucanase of the invention. The specification describes a xyloglucanase enzyme belonging to family 44 of glycosyl hydrolases and exhibiting a relative xyloglucanase activity of at least 30% at pH 5-8. The enzyme exhibits high performance in detergent compositions and prevents binding of certain soils to the xyloglucan left in the cellulosic material. It is stable at pH 5-10 at room temperature and has a half life of more than 50 days when incubated in a full formulated liquid detergent at 30 plus degrees celsius. The enzyme is used in detergent compositions, textile industry for improving the properties of cellulosic fibers, yarn, and woven or non-woven fabrics, preferably in textile scouring process, and in cellulose fiber processing industry for ratting of fibers e.g. hemp, jute, flax, and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              New xyloglucanase enzyme belonging to glycosyl hydrolases family, useful for detergent compositions, and textile or cellulose fiber processing industries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              32.1%; Score 150.5; DB 22; Length 1350; 31.8%; Pred. No. 2.2e-08; Live 21; Mismatches 36; Indels 1;
                                                                                                                                                                                                                                              xyloglucanase; family 44; glycosyl hydrolase; detergent; cellulosic fiber; textile scouring.
                                                                                                                                                                                                                                                                                                                                                           /note= "Glu encoded by GAACCG"
                                                                                                                                                                                                                  Amino acid sequence of xyloglucanase enzyme.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Schuelein M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Page 86-89; 97pp; English.
AAG63963 standard; Protein; 1350 AA.
                                                                                                                                                                                                                                                                                                                            Location/Qualiflers
                                                                                                                                                                                                                                                                                                                                                                                                                                                          21-FEB-2001; 2001WO-DK00116.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24-FEB-2000; 2000DK-0000291.
                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Jorgensen PL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                            Paenibacillus polymyxa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (NOVO ) NOVOZYMES AS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2001-522819/57.
                                                                                                                                                                                                                                                                                                                              Key
Misc-difference 1347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
Matches 27; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1350 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAH75060
                                                                                                                                                                                                                                                                                                                                                                                           WO200162903-A1.
                                                                                                                                                                                 29-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Schnorr K,
                                                                                                                                                                                                                                                                                                                                                                                                                           30-AUG-2001
                                                                                                                                                 AAG63963;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                    AAG63963
                                 qq
```

δy

64 MCCGNIRASFGSVNPATPTADTYLQ 88 : 1 : 1 | 1 | 1: 1258 lscsklngklvkmdkaatgadyyle 1282

g

Search completed: July 2, 2002, 09:15:52 Job time: 240 sec

THIS PAGE BLANK (USPTO)

```
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 47:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 616 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEFAX: 215-540-5818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 <Unknown>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STATE: PA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NUMBER OF
                                                                                                                                                                                                                                                                                                                                                                                                                                         OF.
                                                                                                                                                                                                                                                                                              US-09-136-574A-47
                                                                                                                                       61
61
60.5
60.5
60.5
                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE
                                                                      2, 2002, 09:12:55; Search time 39.87 Seconds (without alignments) 53.912 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Appli
Appl
Appl
Appl
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Appli
Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Appli
Appli
Appli
Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Appli
Appl
Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 43,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 10,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 2,
Sequence 1,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 1
Sequence 7
Sequence 2
Sequence 2
Sequence 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
Sequence
Sequence
                                                                                                                                                   1 VSGGVKVQYKNNDSAPGDNQ.....IRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                             231628
                                                                                                                                                                                                                                                                                                                                                                       /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-033-537A-1
PCT-US95-13813-9
US-09-198-956-10
US-09-198-955A-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3-09-006-632-7
3-09-277-716-22
3-08-330-394A-29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-07-862-588B-7
US-08-048-164A-2
US-08-0460-457-2
US-08-0460-457-2
US-08-460-458-2
US-08-310-358-2
US-08-310-358-2
US-09-006-635-7
US-09-006-632-7
US-09-277-716-22
                                                                                                                                                                                                                                          Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-07-862-588B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -09-118-319-5
-08-353-400-36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-423-439-38
US-08-353-400-33
US-09-423-439-32
                                                                                                                                                                                                                  231628 segs, 24425594 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUMMARIES
                                                                                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                 OM protein - protein search, using sw model
                                                                                                                                                                                       Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                             Issued_Patents_AA:*
                                                                                                                                                                                                                                                                     length: 0
length: 2000000000
                                                                                                                          US-09-917-376-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length
                                                                                                                                                                              BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query
                                                                                                                                                                                                                                                                   Minimum DB seq
Maximum DB seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   120.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           62.5
62.5
62.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          194.5
191.5
172
165.5
165
                                                                                                                          Title:
Perfect score:
                                                                                                                                                                             Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               63.
                                                                                                                                                     Sequence:
                                                                                                                                                                                                                  Searched:
                                                                                                                                                                                                                                                                                                                                                           Database
                                                                         on
::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ş.
```

```
Fabrics Using Truncated
                                         Appl
Appl
Appl
                                                                  App]
                                        sednence
Sednence
Sednence
Sednence
Sednence
Sednence
Sednence
Sednence
Sednence
Sednence
                                                                                                                  Sequence
Sequence
Sequence
                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                              Box
                                                                                                                                                                                                                                                                                            INVENTION: Compositions and Methods for
Treating Cellulose Containing
Cellulase Enzyme Compositions
                                                                                                                                                                                                                                                                                                                                            STREET: Spring House Corporate Center, P.O. CITY: Spring House
      ATORNEY/ACHT INCOMMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: 1997US001/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER: US/09/136,574A FILING DATE: 19-Aug-1998 CLASSIFICATION: <Unknown>
                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/932,571
                                                                                                                                                                    ALIGNMENTS
                                                                                                                                                                                                         Sequence 47, Application US/09136574A.
Patent No. 6294366
GENERAL INFORMATION:
Anderson, Barge
Gibbs, Moreland
Bergquist, Peter
Daniels, Roy
Morgan, Hugh W.
                                                                                                                                                                                                                                                                                                                                     ADDRESSEE: Howson and Howson
                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER: IBM COMPATIBLE OPERATING SYSTEM: DOS CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                              MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                             ZIP: 19477 COMPUTER READABLE FORM
                                                                                                                                                                                                                                                                                                                             CORRESPONDENCE ADDRESS
802
1290
1128
307
307
316
316
316
316
301
617
617
617
617
                                                                                                                                                                                                                                                                                                                      SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                      COUNTRY: USA
133.3
1133.1
1133.1
1133.0
1133.0
1133.0
1133.0
1133.0
1133.0
62.5
61.5
61.5
61.5
61.5
61.5
61.5
```

셤 ò g

```
Daniels, Roy
Morgan, Hugh W.
Williams, Dlane P.
TITLE OF INVENTION: Compositions and Methods for
Treating Cellulose Containing Fabrics Using Truncated
Cellulase Enzyme Compositions
                          4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 1426;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Center, P.O. Box 457
CITY: Spring House
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: 1997US001/CIP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             40.8%; Score 191.5; DB 443.5%; Pred. No. 9.5e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CURRENT APPLICATION DATA:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,574A
FILING DATE: 19-Aug-1998
CLASSIFICATION: CURKNOWN>
PRIOR APPLICATION NUMBER: US 08/932,571
FILING DATE: September 19, 1997
ATTORNEY, AGENT INFORMATION:
NAME: Bak, Mary E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     15; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ) MOLECULE TYPE: NO., 6294366e
; SEQUENCE DESCRIPTION: SEQ ID NO: 43:
US-09-136-574A-43
                                                                                                                                                                                                                                                                                 APPLICANT: Farrington, Graham K.
                                                                                                                                                                                                                     Sequence 43, Application US/09136574A Patent No. 6294366 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 :| |: :| :: | 1 |: | 472 IGASNVTFNFVKLSSGVSGADYYLE 496
                                                                                                               737 IGASNVTFNFVKLSSGVSGADYYLE 761
                                                                                64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH: 1426 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                    Anderson, Paige
Gibbs, Moreland
Bergquist, Peter
                                                                                                                                                                                                                                                                                                                                             Peter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 43:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE: PA
COUNTRY: U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ZIP: 19477
                                                                                                                                                                                  RESULT 3
US-09-136-574A-43
                                         g
                                                                                                                      pp
                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Treating Cellulose Containing Fabrics Using Truncated Cellulase Enzyme Compositions
                                                                                                     1;
                                                                                                     Gaps
                                                                                                                                        1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                 Gaps
                                                                                                     ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 1751;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STREET: Spring House Corporate Center, P.O. Box 457 CITY: Spring House
                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                    Anderson, Paige
Glabs, Moreland
Bergquist, Peter
Daniels, Roy
Morgan, Hugh W.
Williams, Diane P.
TITLE OF INVENTION: Compositions and Methods for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: US 08/932,571
APPLICATION NUMBER: US 08/932,571
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DCKEY NUMBER: 1997US001/CIP
TELECOMMUNICATION INFORMATION:
TELEFRAN: 215-540-9200
TELEFRAX: 215-540-5818
                                                            DB 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  41.5%; Score 194.5; DB 4;
44.7%; Pred. No. 5.7e-13;
Live 14; Mismatches 32;
                                                          ; Score 195.5; DB 4
; Pred. No. 1.2e-13;
15; Mismatches 34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER: IBM Compatible
OPERATING, SYSTEM: DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,574A
FILING DATE: 19-Aug-1998
CLASSIFICATION DATA:
PRIOR APPLICATION OATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-09-136-574A-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
                                                                                                                                                                                                                       61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                              60 WAQIGASNVTFNFVKLSSGVSGADYYLE 87
                                                                                                                                                                                                                                                                                                                                                          Sequence 44, Application US/09136574A
Patent No. 6294366
GENERAL INFORMATION:
APPLICANT: Farrington, Graham K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH: 1751 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 44:
                                                          41.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE CHARACTERISTICS:
                                                                          Best Local Similarity 43.2
Matches 38; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           38; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY: USA
ZIP: 19477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
Matches 38; Conserve
                                                                                                                                                                                                                                                                                                                                     US-09-136-574A-44
US-09-136-574A-47
                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
```

ij

Gaps

```
TITLE OF INVENTION: In Stone Washing
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSE: No. 59580830 No. 5958083disk of No. 5958083th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ramot University Authority for Applied
Research and Industrial Development Ltd.
Technion Research and Development Foundation Ltd.
Bayer, Edward A.
Morag, Ely
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 551;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: Shoham, Yuval
TITLE OF INVENTION: MODIFIED CELLULOSE-BINDING DOMAIN (CBD)
TITLE OF INVENTION: PROTEINS AND USE THEREOF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Yeda Research and Development Co. Ltd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                35.3%; Score 165.5; DB 2;
llarity 37.9%; Pred. No. 2.2e-10;
Conservative 22; Mismatches 29;
                                                                                                                                                                                                                              COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FSSISEO for Windows Version 2.0
CURRENT PEPLICATION DATA:
FILLING DATE: 02-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             E: Browdy and Neimark
419 Seventh Street N.W., Ste. 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE/DOCKET NUMBER: 4492.204-US
TELECOMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
In Stone Washing
                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE: 08-SEP-1995
APPLICATION NUMBER: PCT/DK96/00364
FILING DATE: 03-SEP-1996
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   461 AQIGASNVSAAF -- ANFTGSNTDTYVE 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 9, Application PC/TUS9513813
GENERAL INFORMATION:
APPLICANT: Yeda Research and Devel
                                                                                                                                                                                                                                                                                                                                        NAME: Green, Reza
REGISTRATION NUMBER: 38,475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 551 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Lamed, Raphael
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Morag, Ely
Wilchek, Meir
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                     Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy an
STREET: 419 Seventh
CITY: Washington
                                                                                                                                                                                             COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
Matches 33; Conserv
                                                                                                                                                 COUNTRY: U.S.A. ZIP: 10174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRANDEDNESS:
                                                                                                                                                                                                                   MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
US-09-033-537A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 6
PCT-US95-13813-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT:
                                                                                                                                  STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ñ
                                                                                                                                                                                                                                                       ADDRESSEE: No. 59167960 No. 5916796disk of No. 5916796th America, Inc. STREET: 405 Lexington Avenue, 62nd floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 VSGGVKVQYKNND-SAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNC 59
                                                                                                                                                                                                                 An Enzyme Exhibiting Cellulase Activity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 700;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 36.7%; Score 172; DB 2; Length 70 Best Local Similarity 44.9%; Pred. No. 5.6e-11; Matches 40; Conservative 18; Mismatches 27; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SOFTWARES PATENTIN Release #10, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/862,588B
FILING DATE: 19920727
CLASSIFICATION 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: D6.7/90
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/DK91/00013
FILING DATE: 19-JAN-1991
ATTOREY/AGENT INFORMATION:
ATTOREY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Onishi, Masahiro
APPLICANT: Fich, Merete
APPLICANT: Toft, Annette Hanne
APPLICANT: Toft, Annette Hanne
APPLICANT: Sh lein, Martin
ITLE OF INVENTION: Prevention Of Back-Staining
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              / Lambiris, Elias J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME: Zelson, Steve T. / Lambiris, El
REGISTRATION NUMBER: 30,335 / 33,728
REFERENCE/DOCKET NUMBER: 3425.204-US
TELECOMMINICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ||| :| || :| || :| || :|| 634
                                                                                                                                                                                                                                                                                                                                                                                                                          MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    60 DWAAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                               APPLICANT: Joergensen, Per Linaa APPLICANT: Sch lein, Martin APPLICANT: Hansen, Christian TITLE OF INVENTION: An Enzyme Exh NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 1, Application US/09033537A Patent No. 5958083 GENERAL INFORMATION:
                                                                                   Sequence 2, Application US/07862588B Patent No. 5916796
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEFAX: 212 867 0298 INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; MOLECULE TYPE: protein US-07-862-588B-2
                                                                                                                                                                                                                                                                                                                                                            COUNTRY: U.S.A.
ZIP: 10017
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AMINO ACID
                                                                                                                                                                                                                                                                                                                                        New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              linear
                                                                                                                               GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY:
                                                             US-07-862-588B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-033-537A-1
                                                                                                                                                                                                                                                                                                                                          STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
```

; ;

Gaps

7

```
||| :||:: |:: : | | | :: ||||||::||| :|||::| || 331 VSGNLKVEFYNSNPSDTTNSINPQFKVTNTGSSAIDLSKLTLRYYYTVDGQKDQTFW-CD 389
10; Gaps
                                                                      1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             An Enzyme Exhibiting Cellulase Activity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 493;
Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: Andersen, Lene N.
APPLICANT: Schulein, Martin
APPLICANT: Schulein, Martin
APPLICANT: Bjornvad, Mads E.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Moller, Soren
APPLICANT: Glad, Sanne O. S.
APPLICANT: Gauppinen, Markus S.
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Signal, 
29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     35.2%; Score 165; DB 4; ilarity 36.1%; Pred. No. 2.1e-10; Conservative 23; Mismatches 29
                                                                                                                                                                                                                                                                         390 HAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLE 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    390 HAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADIYLE 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
23; Mismatches
                                                                                                                                                                                                                      61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQ ID NOS: 32
FastSEQ for Windows Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR APPLICATION NUMBER: 1343/97
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR PELICATION NUMBER: 60/067,249
PRIOR FILING DATE: 1997-12-02
PRIOR APPLICATION NUMBER: 60/067,240
PRIOR FILING DATE: 1997-12-02
PRIOR PAPLICATION NUMBER: 60/067,240
PRIOR FILING DATE: 1997-12-02
PRIOR APPLICATION DATE: 1998-05-06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRIOR FILING DATE: 1998-05-06
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 09/184,217
PRIOR FILING DATE: 1998-11-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 7, Application US/07862588B

Patent No. 5916796

GENERAL INFORMATION:

APPLICANT: JOSTGENSEN, PET Linaa

APPLICANT: Sch lein, Martin

APPLICANT: Hansen, Christian

TITLE OF INVENTION: An Enzyme Exhib

NUMBER OF SEQUENCES: 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 12, Application US/09198955A
Patent No. 6187580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM: Clostridium thermocellum
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
Matches 35; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NUMBER OF SEQ ID NOS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-198-955A-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-198-955A-12
35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-07-862-588B-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQ ID NO 12
   Matches
                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 VSGNLKVEFYNSNPSDTTNSINPQFKVTNTGSSAIDLSKLTLRYYYTVDGQKDQTFW-CD 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JOS-198-950-10

Sequence 10, Application US/09198956

Fatent No. 6165769

GENERAL INFORMATION:

APPLICANT: Andersen, Lene N.

APPLICANT: Schulein, Martin

APPLICANT: Lange, Niels Erik K.

APPLICANT: Bjornvad, Mads E.

APPLICANT: Bornorr, Kirk

TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus

TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus

TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus

FILE REFERENCE: 5377.200-US

CURRENT APPLICATION NUMBER: US/09/198,956

CURRENT APPLICATION NUMBER: 1344/97

EARLIER FILING DATE: 1997-11-24

EARLIER FILING DATE: 1997-11-24

EARLIER FILING DATE: 1997-11-24

SEARLIER FILING DATE: 1997-11-02

NUMBER OF SEQ ID NOS: 26

SOFTWARE: FastSEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match . 35.2%; Score 165; DB 5; Length 167; Best Local Similarity 36.1%; Pred. No. 5.7e-11; Matches 35; Conservative 23; Mismatches 29; Indels 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 165; DB 4; Length 493;
Pred. No. 2.1e-10;
                                                                                             COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13813
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64 HAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLE 100
                                                                                                                                                                                                                                                                                                                                                                                                                                            ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: BAYER-3
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM: Bacillus licheniformis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         35.2%;
36.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
TELES: 248633
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               167 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; MOLECULE TYPE: peptide PCT-US95-13813-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-198-956-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-198-956-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

ä

```
TOPOLOGY: linear; MOLECULE TYPE: protein US-08-048-164A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        66141 PENNIE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       New York
: U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CITY: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY: U ZIP: 10036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-460-462-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59167960 No. 5916796disk of No. 5916796th America, Inc.
STREET: 405 Lexington Avenue, 62nd floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25.7%; Score 120.5; DB 2; Length 531; 32.9%; Pred. No. 1.9e-05; Live 18; Mismatches 36; Indels 3
                                                                                                                                                                   MEDIUM TYPE: FIPPPY disk
COMPUTER: IBM PC COMPAtible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/862,588B
FILING DATE: 19920727
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DK 164/90
FILING DATE: 19-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/DK91/00013
FILING DATE: 18-JAN-1991
ATTORNEX/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Shoseyov, oded
APPLICANT: Shpiegi, Itai
APPLICANT: Shpiegi, Itai
APPLICANT: Goldstein, Marc A.
APPLICANT: Doi, Roy H.
TITLE OF INVENTION: CELULOSE BINDING DOMAIN
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME: Zelson, Steve T. / Lambiris, Elias J. REGISTRATION NUMBER: 30,335 / 33,728 REFERENCE/DOCKET NUMBER: 3425.204-US TELECOMUNICATION INFORMATION: TELEPHONE: 212 867 0123 TELEFAX: 212 867 0298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | :| |: :| |: 11|
491 AQIGRTNVLLAF--ANFTGSNTDTY 513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             62 AAMGCGNIRASFGSVNPATPTADTY 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 2, Application US/08048164A Patent No. 5496934
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO: 7:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        : 531 amino acids
AMINO ACID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 32.9%
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS:
                                                                                     STATE: New York
COUNTRY: U.S.A.
ZIP: 10017
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy di
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; MOLECULE TYPE: protein US-07-862-588B-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New York
: U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
                                                                   New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STATE: New COUNTRY: U.SIP: 10036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-048-164A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
```

```
Sequence 2, Application US/08460462
Patent No. 5670623
Patent No. 5670623
GENERAL INFORMATION:
APPLICANT: Shoseyov, Oded
APPLICANT: Shpiegl, Itai
APPLICANT: Goldstein, Marc A.
APPLICANT: OF INVENTION: METHODS OF USE OF CELLULOSE BINDING DOMAIN PROTEINS
ITILE OF INVENTION: METHODS OF USE OF CELLULOSE BINDING SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   24.3%; Score 114; DB 1; Length 162; 29.9%; Pred. No. 2.3e-05; Pred. 22; Mismatches 34; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER TRADABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION UNMER: US/08/460,462
FILING DATE: CONCULTENTLY herewith
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
US/08/048,164A
FILING DATE: 14-APR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           62 AAMGCGN-----IRASF--GSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/048,164
FILING DATE: 14-APR-1993
ATTORNEY/AGENT INFORMATION:
                                                                                             CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7809-(
TELECOMMUNICATION INFORMATION:
TELEFAX: (212) 790-9090
TELEFAX: (212) 790-9090
TELEFAX: (414) PENNIE
TELERX: 66141 PENNIE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (212) 790-9090
(212) 869-8864/9741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE/DOCKET NUMBER: 78( FELECOMMUNICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 29.9%
Matches 29; Conservative
```

```
RESULT 13
US-08-460-458-2
; Sequence 2, Application US/08460458
; Patent No. 573694
; GENERAL INFORMATION:
TITLE OF INVENTION: KITS AND METHODS OF DETECTION USING CELLULOSE BINDING DOMAI
TITLE OF INVENTION:
NUMBER OF SEQUENCES:
CORRESPONDENCE ADDRESS:
ADDRESSEE:
A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3 TSSMSVEFYNSNKSAQTNSITPIIKIINTSDSDLNLNDVKVRYYYYTSDGTQGQTFW-CDH 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
              2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                     3 TSSMSVEFYNSNKSAQTNSITPIIKITNTSDSDLNLNDVKVRYYYTSDGTQGQTFW-CDH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 162;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
SOFTWARE: Patentin Release #1.0, Version #1.25
CURREWT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,458
FILING DATE: concurrently herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/048,164
FILING DATE: 14-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REFERENCE/DOCKET NUMBER: 7809-007
TELECOMMUNICATION: INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 24.3%; Score 114; DB 1; Best Local Similarity 29.9%; Pred. No. 2.3e-05; Matches 29; Conservative 22; Mismatches 34
                                                                                                                                                                                                62 AAMGCGN-----IRASF--GSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                     62 AGALLGNSYVDNTSKVTANFVKETASP-TSTYDTYVE 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62 AAMGCGN-----IRASF--GSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62 AGALLGNSYVDNTSKVTANFVKETASP-TSTYDTYVE 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
CINTRY: U.S.A.
ZIP: 10036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 2, Application US/08460455; Patent No. 583/814
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Shoseyov, Oded
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    162 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; MOLECULE TYPE: protein US-08-460-458-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 14
US-08-460-455-2
                                                                                                g
                                                                                                                                                                                                δy
                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: Shoseyov, oded
APPLICANT: Shoseyov, oded
APPLICANT: Shpiegl, Ital
APPLICANT: Shpiegl, Ital
APPLICANT: Sholegl, Ital
APPLICANT: Soldstein, Marc A.
APPLICANT: Doi, ROY H.
TITLE OF INVENTION: CELLULOSE BINDING DOMAIN FUSION PROTEINS
NUMBER OF SEQUENCES: 21
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 24.3%; Score 114; DB 1; Length 162; Best Local Similarity 29.9%; Pred. No. 2.3e-05; Matches 29; Conservative 22; Mismatches 34; Indels 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   24.3%; Score 114; DB 1; Length 162; 29.9%; Pred. No. 2.3e-05; Live 22; Mismatches 34; Indels 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     34; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER EADDABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Floppy 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 2, Application US/08460457 Patent No. 5719044 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (212) 790-9090
(212) 869-8864/9741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   : 162 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 24.3
Best Local Similarity 29.9
Matches 29; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                             ; MOLECULE TYPE: protein US-08-460-462-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; MOLECULE TYPE: protein US-08-460-457-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               New York
: New York
RY: U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CITY: N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
```

4

```
Search completed: July
Job time: 64 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 ISSMSVEFYNSNKSAQTUSITPIIKITNISDSDLNLNDVKVRYYYTSDGTQCQTFW-CDH 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 24.3%; Score 114; DB 2; Length 162; Best Local Similarity 29.9%; Pred. No. 2.3e-05; Matches 29; Conservative 22; Mismatches 34; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 2, Application US/08330394A
Patent No. 5856201
GENERAL INFORMATION
APPLICANT: Shoseyov, Oded
APPLICANT: Shoseyov, Oded
APPLICANT: Shosejov, Ital
APPLICANT: Oddstein, Marc A.
TITLE OF INVENTION: BETHODS OF DETECTION USING THE
TITLE OF INVENTION: METHODS OF DETECTION USING THE
TITLE OF INVENTION: CELLULOSE BINDING DOMAIN
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSED: PENNIE & EDWONDS
STREET: 1155 AVENUE OF THE AMERICAS
APPLICANT: Goldstein, Marc A. APPLICANT: Dol, Roy H. TITLE OF INVERVION: CELLULOSE BINDING DOMAIN PROTEINS NUMBER OF SEQUENCES: 21
                                                                                                                                                                                                                                                                                                                                                                            SOFTWARE: Patentin Release #1.0, Version #1.25 SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,455 FILING DATE: Concurrently herewith CLASSIFICATION: 435 PRIOR APPLICATION DATA: APPLICATION NUMBER: US 08/048,164 FILING DATE: 14-APR-1993 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62 AAMGCGN-----IRASF--GSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62 AGALLGNSYVDNTSKVTANFVKETASP-TSTYDTYVE 97
                                                                                               CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7809-009
                                                                                                                                                                                                                                                  ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME: MISTOCK, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7809-(
TELECOMUNICATION INFORMATION:
TELEPAN: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       : 162 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; MOLECULE TYPE: protein US-08-460-455-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STATE: NEW YORK COUNTRY: U.S.A. ZIP: 10036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CITY: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-330-394A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QQ
```

```
2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 162;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        34; Indels
                                   COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOCTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/330,394A
FILING DATE: 27-OCT-1994
PRIOR APPLICATION NUMBER: CIP OF PCT/US94/04132
FILING DATE: 14-APR-1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 114; DB 2;
Pred. No. 2.3e-05;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                62 AAMGCGN-----IRASF--GSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2, 2002, 09:12:56
                                                                                                                                                                                                                                  NAME: MISROCK, S. LESLIE
REFERENCE/DOCKET NUMBER: 7809
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELERAX: (212) 869-8864
TELEX: 66441 PENNIE
INFORMATION FOR SEQ ID NO: 2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24.3%; 29.9%;
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH: 162 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 29.9%
Matches 29; Conservative
                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; MOLECULE TYPE: protein US-08-330-394A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY:
```

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

July Run on:

2, 2002, 09:14:14; Search time 54.67 Seconds (without alignments) 154.671 Million cell updates/sec

Perfect score:

US-09-917-376-5 469 1 VSGGVKVQYKNNDSAPGDNQ.....IRASFGSVNPATPTADTYLQ Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

283138 segs, 96089334 residues Searched:

283138 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

pir1:\* pir2:\* pir3:\* pir4:\* Database

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## STIMMARTES

|        |       | æ     |        |     | SUMMAKIES |                      |
|--------|-------|-------|--------|-----|-----------|----------------------|
| Result |       | Query |        |     |           |                      |
| . ON . | Score | Match | Length | DB: | ID        | Description          |
| 1      | 191   | 40.7  | 1711   | 7   | T31337    | . 1,4-beta-qlucanase |
| 7      | 190.5 | 40.6  | 1779   | 7   | T31085    | xylanase - Caldice   |
| 3      | 190   | 40.5  |        | 7   | G69593    | e (EC 3              |
| 4      | 190   | 40.5  |        | ~   | A26874    | cellulase (EC 3.2.   |
| 5      | 188   | 40.1  |        | 7   | JN0111    | e (EC 3              |
| 9      | 186.5 | 39.8  | 1331   | 7   | A48954    | ndo-1,4              |
| 7      | 186.5 | 39.8  | 1742   | ~   | T17120    | cellulase (EC 3.2.   |
| 89     | 186   | 39.7  | 499    | 7   | A27198    | e e                  |
| 6      | 182.5 | 38.9  | 915    | ~   | A43802    | cellulase (EC 3.2.   |
| 10     | 182.5 | 38.9  | 1039   | ~   | 802711    | cellulase (EC 3.2.   |
| 11     | 180.5 | 38.5  | 145    | 7   | A41897    | ğ                    |
| 12     | 180   | 38.4  | 486    | 7   | I40548    | bifunctional cellu   |
| 13     | 172   | 36.7  | 700    | ~   | B41897    | cellulase (EC 3.2.   |
| 14     | 165.5 | 35.3  | 879    | ~   | A47704    | endoglucanase I (E   |
| 15     | 165   | 35.2  | 1854   | ~   | S36859    | cipA protein - Clo   |
| 16     | 149.5 | 31.9  | 505    | ~   | S39962    | endoglucanase - Er   |
| 17     | 147.5 | 31.4  | 504    | 7   | S54744    | cellulase (EC 3.2.   |
| 18     | 141.5 | 30.2  | 986    | 7   | S12021    | thermoactive cellu   |
| 19     | 130.5 | 27.8  | 586    | 7   | PC6006    | scaffolding protei   |
| 20     | 130.5 | 27.8  | 1483   | ~   | C97012    | probably celluloso   |
| 21     | 126.5 | 27.0  | 1162   | ~   | T30433    | scaffolding protei   |
| 22     | 117   | 24.9  | 1230   | 7   | S47466    | cellulose 1,4-beta   |
| 23     | 114   | 24.3  | 1848   | 7   | A44140    | ρ                    |
| 24     | 75.5  | 16.1  | 618    | 7   | T08685    | hypothetical prote   |
| 25     | 73.5  | 15.7  | 547    | 7   | T25478    | _                    |
| 26     | 73.5  | 15.7  | 1428   | 7   | AC2224    |                      |
| 27     | 70    | 14.9  | 5188   | 7   | B85547    |                      |
| 28     | 70    | 14.9  | 5291   | 7   | F90696    | thetical             |
| 29     | 98.5  | 14.6  | 574    | ~   | A69196    |                      |

| hypothetical prote | thaumatin-like pro | hypothetical prote | ser-lys rich hypot | polyprotein - hepa | probable purH prot | probable reverse g | hypothetical prote | hensin - rabbit | hemolysin - Aquife | sanko - human | endo-beta-1,3-1,4- | replicative DNA he | protein-tyrosine-p | hypothetical prote | Ig heavy chain V r |
|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|-----------------|--------------------|---------------|--------------------|--------------------|--------------------|--------------------|--------------------|
|                    |                    |                    |                    |                    |                    |                    |                    |                 |                    |               |                    |                    |                    |                    | •                  |
| н72684             | JC7201             | T47790             | T40462             | T01075             | C70717             | G72614             | 820669             | T30549          | H70393             | A59386        | H95976             | AB2512             | B44390             | AH2515             | E37267             |
| 7                  | ~                  | 7                  | 7                  | 7                  | 7                  | 7                  | 7                  | 7               | 7                  | 7             | 7                  | 7                  | Н                  | 7                  | 7                  |
| 271                | 247                | 535                | 751                | 1435               | 523                | 1222               | 1571               | 1594            | 398                | 2403          | 269                | 454                | 802                | 4936               | 122                |
| 14.3               | 14.2               | 14.1               | 14.0               | 14.0               | 13.9               | 13.6               | 13.6               | 13.6            | 13.5               | 13.5          | 13.3               | 13.3               | 13.3               | 13.3               | 13.2               |
| 29                 | 66.5               | 99                 | 65.5               | 65.5               | 65                 | 64                 | 64                 | 64              | 63.5               | 63.5          | 62.5               | 62.5               | 62.5               | 62.5               | 62                 |
| 30                 | 31                 | 32                 | 33                 | 34                 | 32                 | 36                 | 37                 | 38              | 39                 | 40            | 41                 | 42                 | 43                 | 44                 | 45                 |

## ALIGNMENTS

| - |          |
|---|----------|
| _ | 7        |
| _ | 3        |
| _ | r        |
| Ω | $\vdash$ |

1,4-beta-glucanase (EC 3.2.1.-) - Anaerocellum thermophilum (fragment)
(;Species: Anaerocellum thermophilum
(;Date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 02-Sep-2000
(;Accession: T31337
R;Zverlov, V.; Mahr, S.; Riedel, K.; Bronnenmeier, K.
Microbiology 144, 457-465, 1998
A;Title: Properties and gene structure of a bifunctional cellulolytic enzyme (CelA) formains.

A. Reference number: 221003; MUID:98154434
A. Reference number: 221003; MUID:98154434
A. Accession: T3137
A. Status: preliminary; translated from GB/EWBL/DDBJ
A. Rotecule type: DNA
A. Molecule type: DNA
A. Rotecule type: DNA
A. Rotecule type: DNA
A. Rotecule type: NA
A. Rotecule ty

Gaps 5; Query Match

40.7%; Score 191; DB 2; Length 1711;
Best Local Similarity 43.8%; Pred. No. 1.2e-11;
Matches 39; Conservative 18; Mismatches 30; Indels

; ;

1 VSGG-VKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNC 59

g

60 DWAAMGCGNIRASFGSVNPATPTADTYLQ 88 δλ

742 DWAQIGASNVTFKFVKLSSSVSGADYYLE 770 g

Yglanase - Caldicellulosiruptor sp.

Yglanase - Caldicellulosiruptor sp.
C;Species: Caldicellulosiruptor sp.
C;Species: Caldicellulosiruptor sp.
C;Species: O2-8ep-2000 #sequence\_revision 02-Sep-2000 #text\_change 02-Sep-2000
C;Accession: T31085
R;Morris, D.D.; Gibbs, M.D.; Ford, M.; Thomas, J.; Bergquist, P.L.
S;Morris, D.D.; Gibbs, M.D.; Ford, M.; Thomas, J.; Bergquist, P.L.
A;Description: Family 10 and 11 xylanase genes from Caldicellulosiruptor sp. Rt69B.1.
A;Reference number: 220972
A;Reference number: 220972
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Residues: 1-1779 <AMRN>
A;Residues: 1-1779 <AMRN>
A;Residues: 1-1779 CAMRN>
A;Canetics: A;Note: xynC

```
A;Cross-references: EMBL:229076; NID:9509266; PIDN:CAA82317.1; PID:9509267
R;Seo. Y. S.; Lee, Y. H.; Pek, U. H.; Kang, H.
Korean J. Microbiol. 24, 236-242, 1986
A;Title: Analysis on the nucleotide sequence of the signal region of Bacillus subitil
A;Reference number: I39803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA 62
A; Reference number: S49103
                                                                                                                               A; Molecule type: DNA
A; Residues: 10-508 <WOL>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Gene: bglC
C; Function:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      cellulase (EC 3.2.1.4) bglC precursor - Bacillus subtilis
Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cispecias: Cisp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Accession: A26114
A; Molecule type: DNA
A; Recession: A26114
A; Molecule type: DNA
A; Recession: A26114
A; Molecule type: DNA
A; Resterences: GB:X04689; NID:939823; PIDN:CAA28392.1; PID:939824
A; Experimental source: strain PAP115
A; Note: part of this sequence, including the amino end of the mature form, was confirmed a source: part of this sequence, including the amino end of the mature form, was confirmed R; Lindahl, V; Aa, K.; Tronsmo, A.
A; Tronsmo an endo-beta-1, 4-glucanase gene from Bacillus subtilis C
A; Recession: 140353; MUID:95225656
A; Accession: 140353; MUID:95225656
A; Accession: 140353; MUID:95225656
A; Molecule type: DNA
A; Relatences: EMBL:X67044; NID:939776; PIDN:CAA47429.1; PID:939777
B; Lindahl, V; Aa, K.
Submitted to the EMBL Data Library, June 1992
A; Reference number: $24239
A; Rocession: $24239
A; Molecule type: DNA
A; Reference number: $24239
A; Molecule type: DNA
A; Rocession: $24239
A; Molecule type: DNA
A; Mol
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      **Residues: 10-291,'N', 293-508 <LINZ>
**A; Residues: EMBL:X67044; NID:g39776; PIDN:CAA47429.1; PID:g39777
**A; Experimental Source: Strain CK-2
**R; Wolf, M.; Geczi, A.; Borriss, R.
**Submitted to the EMBL Data Library, December 1993
**Description: Genes encoding beta glucan-hydrolyzing enzymes in Bacillus subtilis: cons
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Status: nucleic acid sequence not shown; translation not shown
A; Molecule type: DNA
A; Molecule type: DNA
A; Molecule type: DNA
A; Mostacia (B: 299113; GB: AL009126; NID: 92634090; PIDN: CAB13696.1; PID: e1183471;
A; Experimental source: strain 168
A; Exper
                                                                                                                                                                                                          ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                         GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                   Length 1779;
                                                                                                                                                                                                          Indels
       Score 190.5; DB 2;
Pred. No. 1.4e-11;
                                                                                                                                                                                                  14; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1172 IGASNVTFNFVKLTSGVSGADYYLE 1196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                    40.68;
                                                                                                                                                                                                  Conservative
                                                                                                                               Similarity
                                                                                                                                                                                                  37;
                                                    Query Match
Best Local S
Matches 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
A,Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as A;Pathway: cellulose degradation c,Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation F;1-38/Domain: signal sequence #status predicted <SIG>F;39-508/Product: cellulase #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as A;Pathway: cellulose degradation
C;Keywords: alternative initiators; extracellular protein; glycosidase; hydrolase; po
F;1-38/Domain: (or 2-38) signal sequence #status predicted <SIG>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Experimental source: strain DLG
A; Accession: B26874
A; Accession: B26874
A; Molecule type: protein
A; Estdues: 39-53
A; Experimental source: strain DLG
A; Note: the authors believe Met-1 and Met-2 may be alternate initiators
C; Comment: The low molecular weight of the mature protein suggests carboxyl-terminal
C; Function:
                                                                         A;Molecule type: DNA
A;Residues: 10-13,'V',15,'N',17-18,'V',20-21,'F',23,'A',25-26,'AI',29-31,'P',33,'PQ'
A;Cross-references: GB:M38634; NID:g142657; PIDN:AAA22300.1; PID:g142658
A;Experimental source: strain ATCC 6633
C;Comment: The low molecular weight of the mature protein suggests carboxyl-terminal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NiAlternate names: endo-1,4-beta-glucanase
C;Species: Bacillus subtilis
C;Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 15-Oct-1999
C;Accession: A26874, B26874
R;Robson, L.M.; Chambliss, G.H.
J; Bacteriol. 169, 2017-2025, 1987
A;Fille: Endo-beta-1-4-glucanase gene of Bacillus subtilis DLG.
A;Reference number: A26874; MUID:87194581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Cross-references: GB:M16185; NID:g143007; PIDN:AAA22496.1; PID:g143008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             cellulase (EC 3.2.1.4) precursor - Bacillus subtilis (strain DLG)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 508;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 508;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     40.5%; Score 190; DB 2; L
40.7%; Pred. No. 4.3e-12;
tive 21; Mismatches 26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
40.5%; Score 190; DB 2; 1
Best Local Similarity 43.0%; Pred. No. 4.3e-12;
Matches 37; Conservative 17; Mismatches 28;
A; Accession: 139803
A; Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                :||||: | ::: |||||: | 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: DNA
A; Residues: 1-508 <ROB1>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
Matches 35; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Accession: A26874
```

; ;

Gaps

..

Indels

l .

ï

용

g ò

```
cellulase (EC 3.2.1.*) precursor, thermoactive - Caldocellum saccharolyticum cispaces: Caldocellum saccharolyticum cispaces: Caldocellum saccharolyticum C; Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Nov-2000 C; Accession: T17120; A43745
R; Accession: T17120; A43745
R; Title: CelA, another gene coding for a multidomain cellulase from the extreme therm A; Reference number: 218698; MUID:95336703
A; Reference number: 218698; MUID:95336703
A; Accession: T17120
A; Accession: T17120
A; Residues: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 1-1742 < PED>
A; Residues: 1-1743 < PED>
A; Reference number: A43745; MUID:91247819
A; Accession: A3745; MUID:91247819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Cellulase (EC 3.2.1.4) precursor - Bacillus subtilis (strain IFO3034)
N;Alternate names: endo-1,4-beta-glucanase
C;Species: Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 15-Oct-1999
C;Accession: A27198
R;Nakamura, A.; Uozumi, T.; Beppu, T.
Bur. J. Biochem. 164, 317-320, 1987
A;Title: Nucleotide sequence of a cellulase gene of Bacillus subtilis.
A;Reference number: A27198; MUID:87190397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Molecule type: DNA
A;Residues: 1516-1544,'A',1546-1742 <LUE>
A;Cross-references: EMBL:M36063; NID:g144292; PIDN:AAA72860.1; PID:g144293
                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Molecule type: DNA
A;Residues: 1-499 <NAK>
A;Cross-references: GB:M28332; NID:g142670; PIDN:AAA22307.1; PID:g142671
                                                                                                                                                                                                                                                2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                           1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 1742;
                                                                                                      Length 1331;
   C;Keywords: glycosidase; hydrolase; polysaccharide degradation
                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 39.8%; Score 186.5; DB 2; 42.5%; Pred. No. 3.7e-11; iive 18; Mismatches 31;
                                                                                                      Query Match 39.8%; Score 186.5; DB 2; Best Local Similarity 42.5%; Pred. No. 2.8e-11; Matches 37; Conservative 18; Mismatches 31;
                                                                                                                                                                                                                                                                                                                                                                                                                                   763 AQIGASNVTFKFVKLSSSVSGADYYLE 789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                           AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Gene: celA
C;Keywords: glycosidase; hydrolase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local S
Matches 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                           62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Accession: A48954
A; Status: preliminary
A; Status: preliminary
A; Molecule type: nucleic acid
A; Residues: 1-1331 cGIB>
A; CGIB>
A; Cross-references: GB:L01257; NID:g144290; PIDN:AAA71887.1; PID:g144291
A; Cross-reference extracted from NCBI backbone (NCBIN:121576, NCBIP:121577)
B; Lucthi, E.; Bhana Jasmat, N.; Grayling, R.A.; Love, D.R.; Bergquist, P.L.
Appl. Environ. Microbiol. 57, 694-700, 1991
A; Title: Cloning, sequence analysis, and expression in Escherichia coli of a gene coding A; Reference number: A43745; MUID:91247819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: DNA
A; Residues: 1-499 cARR>
A; Residues: 1-499 cARR>
A; Cross-references: GB:D01057; NID:g216387; PIDN:BAA00859.1; PID:d1001323; PID:g216388
A; Note: the authors translated the codon ATA for residue 102 as Tyr
C; Comment: The low molecular weight of the mature protein suggests carboxyl-terminal pro
                                                                                                                                                                                                                                                                                                                                            Cellulase (EC 3.2.1.4) precursor - Bacillus subtilis (strain BSE616)

N;Alternate names: endo-1.4-beta-glucanase
C;Species: Bacillus subtilis
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 28-May-1999
C;Accession: JN0111
A;Park, S.H.; Kim, H.K.; Pack, M.Y.
Agric. Biol. Chem. 55, 441-448, 1991
A;Title: Chem. 55, 441-448, 1991
A;Title: Characterization and structure of the cellulase gene of Bacillus subtilis BSE61
A;Reference number: JN0111; MUID:91299280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as ce
A;Pathway: cellulose degradation
C;Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
F;1-29/Domain: signal sequence #status predicted <SIG>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C;Accession: A48954; B43745
R;Glbbs, M.D.; Saul, D.J.; Luthi, E.; Bergquist, P.L.
R;Glbbs, M.D.; Saul, D.J.; Luthi, E.; Bergquist, P.L.
A;Title: The beta-mannanase from "Caldocellum saccharolyticum" is part of a multidomain A;Reference number: A48954; MUID:93119139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      mannan endo-1,4-beta-mannosidase (EC 3.2.1.78) - Caldocellum saccharolyticum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N.Alternate names: beta-mannanase
C;Species: Caldocellum saccharolyticum
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Residues: 1.1337, PPROHOHRQ' <LUE>
A; Cross-references: EMBL:M36063; NID:g144292; PIDN:AAA72861.1; PID:g144294
A; Note: the authors translated the codon CAC for residue 262 as Glu
A; Note: this sequence has been revised in reference A48954
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        40.1%; Score 188; DB 2; Length 499; 40.0%; Pred. No. 6.7e-12; tive 18; Mismatches 31; Indels
                                                                                                                                                                       419 OMGCGNLTHKFVTLHKPKOGADTYLE 444
                                                                                                      88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LGCGNVTYKFVTLHKPKQGADTYLE 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                      63 AMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
Matches 34; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Accession: JN0111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              : Function:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           411
```

ò Db ò g ï

```
C;Genetics:
A;Gene: celB
C;Function: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as A;Punction: hydrolysis of 1,4-beta-P-glucosidic linkages in beta-D-glucans such as A;Puthway: cellulose degradation
A;Puthway: cellulose degradation
C;Superfamily: Streptomyces endo-1,4-beta-xylanase A homology
C;Superfamily: Streptomyces endo-1,4-beta-xylanase A homology csxx>
F;29-1039/Product: cellulase #status predicted <MAT>
F;20-1039/Product: cellulase #status predicted <MAT>
           A; Cross-references: EMBL: X13602; NID: 940645; PIDN: CAA31936.1; PID: 940646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               : | |: | : : | |: | 479 QIGASNVTFKFVKLSSSVSGADYYLE 504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  63 AMCCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  38.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         38.9%;
ilarity 41.9%;
Conservative 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           67 GNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SNIQISEG--NHTGTNSDTYVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 36; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Accession: I40548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Simi
Matches 36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  cellulase (EC 3.2.1.4) / cellulose 1,4-beta-cellobiosidase (EC 3.2.1.91) - Caldocellum s Caldocellum seconario and can be caldocellum saccharolyticum (c) species: Caldocellum (c) species: Cal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                tra
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ex
O
A;Experimental source: strain IFO3034
C;Function:
A;Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as A;Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as A;Pathway: cellulose degradation
C;Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
F;1-36/Domain: signal sequence #status predicted <SIG>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Callulase (EC 3.2.1.4) precursor - Caldocellum saccharolyticum
N;Alternate names: endo-1,4-beta-glucanase
N;Contains: callulase (EC 3.2.1.4); cellulose 1,4-beta-cellobiosidase (EC 3.2.1.91)
C;Species: Caldocellum saccharolyticum 07-Sep-1990 #text_change 08-Oct-1999
C;Bete: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Oct-1999
C;Accession: S02711
R;Saul. D.J.; Williams, L.C.; Love, D.R.; Chamley, L.W.; Bèrgquist, P.L.
Nucleic Acids Res. 17, 439, 1989
A;Title: Nucleotide sequence of a gene from Caldocellum saccharolyticum encoding for A;Reference number: S02711; WUID:89098398
A;Accession: S02711
A;Molecule type: DNA
A;Residues: 1-1039 <SAU>
                                                                                                                                                                                                                                                                                                                                                                      ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 915;
                                                                                                                                                                                                                                                                                            Length 499;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         31;
                                                                                                                                                                                                                                                                                        39.7%; Score 186; DB 2; L 40.7%; Pred. No. 1.1e-11; ive 19; Mismatches 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               38.9%; Score 182.5; DB 2.41.9%; Pred. No. 4.8e-11; ive 18; Mismatches 31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          :||||: | ::: | ||||: | 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OIGASNVTFKFVKLSSSVSGADYYLE 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         36; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
Matches 36; Conserv
                                                                                                                                                                                                                                                                                                                             Local Similarity
nes 35; Conserv
                                                                                                                                                                                                                                                                                            Query Match
Best Local S:
Matches 35,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
```

```
C;Species: Bacillus lautus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 15-Oct-1999
C;Datesslon: A41897; S27498
R;Hansen, C.K.; Didexichsen, B.; Jorgensen, P.L.
J. Bacteriol. 174, 3522-3531, 1992
A;Title: celA from Bacillus lautus PL236 encodes a novel cellulose-binding endo-beta-A;Reference number: A41897; MUID:92276330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               and its structural gene: the
                                                              ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .
.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C;Species: Bacillus sp. C;Accassion: 12-Aug-1996 #text_change 15-Oct-1999 C;Bate: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 15-Oct-1999 C;Accassion: 140648 R;Han, S.J.; Yoo, Y.J.; Kang, H.S. J. Blol. Cham. 270, 26012-28019, 1995 J;Blol. Cham. 270, 26012-28019, 1995 A;Fitle: Characterization of a bifunctional cellulase and its structural get A;Reference number: 140548; MUID:96029707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Molecule type: DNA
A; Residues: 1-145 <han>
A; Cross-references: EMBL:M76588; NID:g142661; PIDN:AAA22302.1; PID:g142662
A; Experimental source: PL236
A; Note: sequence extracted from NCBI backhone (NCRTD:104604)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molacule type: DDA
A;Residues; 1-486 <RES-
A;Cross-references: EMBL:U27084; NID:g857575; PIDN:AAC43478.1; PID:g857576
                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                          GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           7 VQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMGC 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   <u>ښ</u>
   Length 1039;
                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Experimental source: PL236
Note: sequence extracted from NCBI backbone (NCBIP:104604)
; Score 182.5; DB 2;
; Pred. No. 5.6e-11;
18; Mismatches 31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 180.5; DB 2;
Pred. No. 1.1e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               bifunctional cellulase precursor - Bacillus sp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                          cellulase homolog - Bacillus lautus (fragment)
```

```
A; Status: preliminary
A; Molecule type: DNA
A; Molecule type: DNA
A; Molecule type: DNA
A; Residues: 1-1884 <GER>
A; Residues: 1-1884 <GER>
A; Residues: 0.T.; Romaniec, M.P. M.; Kobayashi, T.; Huskisson, N.S.; Demain, A.L.
Mol. Microbiol. 8, 325-334, 1993
A; Title: Sequencing of a Clostridium thermocellum gene (cipA) encoding the cellulosom
A; Reference number: S33527; MUID:93302508
A; Accession: S33527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Status: preliminary
A; Molecule type: DNA
A; Molecule type: DNA
A; Residues: 1-1615,1617-1854 <GE2>
A; Residues: 1-1615,1617-1854 <GE2>
A; Cross-references: EMBL:L08665
B; Fujino, T.; Beguin, P.; Aubert, J.P.
FEMS Microbiol. Lett. 94, 165-170, 1992
A; Title: Cloning of a Clostridium thermocellum DNA fragment encoding polypeptides tha A; Reference number: S25767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Cross-references: EMBL:X67406
R; Beguin, P.
submitted to the EMBL Data Library, August 1992
A; Reference number: $28659
A; Molecule type: DNA
A; Residues: 1216-1232, Y', 1234-1241, 'T', 1243-1319, 'R', 1321-1615, 1617-1768, 'R', 1770-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Cross-references: EMBL:X67406
K;Fujino, T.; Beguin, P.; Aubert, J.P.
Bacteriol. 175, 1891-1899, 1993
A;Title: Organization of a Clostridium thermocellum gene cluster encoding the cellulo
                                                                      ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Accession: S25767
A; Molecule type: DNA
A; Residues: 1216-1232, Y',1234-1241, T',1243-1319, R',1321-1615,1617-1854 <FUJ>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CipA protein - Clostridium thermocellum
N;Alternate names: probable cellulosome protein large chain SL
C;Species: Clostridium thermocellum
C;Species: Clostridium thermocellum
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Dec-2000
C;Accession: S36859; S33527; S25767; S28659; T18261
R;Gerngross, U.T., Demain, A.L.
A;Reference number: S36859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Cross-references: EMBL:X67506; NID:g296879; PID:g296880; PIDN:CAA47840.1
C;Genetics:
A;Gene: cipA
                                                                                                                                                                         365 VSGNLKVEFYNSNPSDTTNSINPQFKVTNTGSSAIDLSKLTLRYYYTVDGQKDQTFW-CD 423
                                                                          Gaps
                                                                                                                                  1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10;
                                                                          Ĥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               35.2%; Score 165; DB 2; Length 1854; 36.1%; Pred. No. 6.9e-09; ive 23; Mismatches 29; Indels 1
       Length
                                                                          Indels
   DB 2;
                                                                      32;
; Score 165.5; DB 2; Pred. No. 2.7e-09; 26; Mismatches 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                 795 WSSAGNSNVTGNFFNLSSPKEGADTCLE 822
                                                                                                                                                                                                                                                                            61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Reference number: Z18847; MUID:93209931
   35.3%;
33.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity 36.1
                                                                      29; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             <FU2>
   Query Match
Best Local Similarity
Matches 29; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: DNA
A; Residues: 1821-1854
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Accession: S36859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Accession:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                            δλ
                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        cellulase (EC 3.2.1.4) - Bacillus lautus
N;Alternate names: endo-1,4-beta-glucanase
C;Specias: Bacillus lautus
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 15-Oct-1999
C;Accession: B41897; S27499
R;Hansen, C.K.; Diderichsen, B.; Jorgensen, P.L.
J. Bacteriol. 174, 3522-3531, 1992
A;Title: cellA from Bacillus lautus PL236 encodes a novel cellulose-binding endo-beta-1,4
A;Reference number: A41897; MUID:92276330
A;Retus: preliminary
A;Molecule type: DNA
A;Retaus: 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              c. Species: Clostridium thermocellum
C; Species: Clostridium thermocellum
C; Species: Clostridium thermocellum
C; Species: 19-Dec-1993 #sequence_revision 18 Nov-1994 #text_change 15-Oct-1999
C; Accession: A47704 6.P.; Davidson, K.; Laurie, J.I.; Huskisson, N.S.; Gilbert, H.J.
J. Gen. Microbiol. 139, 307-316, 1993
A. Title: Gene sequence and properties of Cell, a family E endoglucanase from Clostridium A; Reference number: A47704 MuID:93171873
A; Reference preliminary A; Molecule type: DNA; protein
A; Residues: 1-879 <4A2>
                                                                                                                                                                                                         ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        'n
                                                                                                                                                                                                                                                                                                               A;Cross-references: GB:L04735; NID:9144807; PIDN:AAA20892.1; PID:9144808
A;Note: sequence extracted from NCBI backbone (NCBIN:125637, NCBIP:125638)
C;Keywords: glycosidase; hydrolase
                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                        4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 VSGGVKVQYKNND-SAPGDNQIKPGLQLVNTGSSSSVDLSTVTVRYWFTRDGGSSTLVYNC 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4;
                                                                                                                                                                                                         5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 700;
                                                                                                                                      Length 486,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    27; Indels
                                                                                                                                                                                                         32;
                                                                                                                                  ; DB 2;
4.5e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 172; DB 2;
Pred. No. 4.5e-10;
                                                                                                                               tch 38.4%; Score 180; DB al Similarity 38.8%; Pred. No. 4.5e-33; Conservative 18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 60 DWAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                  + 11 LGCGNVSHTVVTLHKPKQGADTYLE 435
                                                                                                                                                                                                                                                                                                                                                                                                        64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match 36.7%;
Best Local Similarity 44.9%;
Matches 40; Conservative 16
                                                                                                                                                                  Best Local Similarity
Matches 33; Conserv
                                                                                                                                  Query Match
C;Genetics:
A;Gene: cel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 14
                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                        台
                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
```

Search completed: July 2, 2002, 09:14:14 Job time: 142 sec

us-09-917-376-5.rsp

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                       bacillus th
aleuria aur
caenorhabdi
bacillus su
bordetella
                                                                                                                                                                    microbispor
mus musculu
                     escherichia
                                           escherichia
                                                                                                                                                      homo sapien
                                                                                                                                                                                                            aspergillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-DLG;
MEDLINE-87104581; PubMed-3106328;
MEDLINE-87104581; PubMed-3106328;
Rebsson L.M., Chambliss G.H.;
"Endo-beta-1,4-glucanase gene of Bacillus subtilis DLG.";
J. Bacteriol. 169:2017-2025(1987).
-!- CATALYITIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
linkages in cellulose.
-!- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PROTON DONOR (BY SIMILARITY).
UNCLEOPHILE (BY SIMILARITY).
CELLULOSE-BINDING (BY SIMILARITY).
333904EE95A63EE1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1991 (Rel. 20, Last sequence update)
01-OCT-1994 (Rel. 30, Last annotation update)
Endoglucanase precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4;
                 P46849
P58127
Q45733
P18891
P34446
P17740
O75698
P075698
P075698
P075698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                40.5%; Score 190; DB 1; Length 499;
43.0%; Pred. No. 4.4e-14;
tive 17; Mismatches 28; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PIR; A26874; A26874.
HSSP: O85465; 1A3H.
InterPro; IPR001956; GBD_3.
Pfam; PF00942; CBD_3; 1.
Pfam; PF00150; cellulase; 1.
PROSTIE; PS00659; GLYCOSYL_HYDROL_F5; 1.
PROSTIE; PS00659; GLYCOSYL_HYDROL_F5; 1.
PROSTIE; PS00659; GLYCOSYL_HYDROL_F5; 1.
29; Hydrolase; Glycosidase; Signal.
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
                                                                                                                                                                                                                                                                                                                                                            499 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ENDOGLUCANASE.
                                                                                                                                                                                                                                                                 ALIGNMENTS
HV50_MOUSE
RTCA_ECOLI
RTCA_ECO57
C9CA_BACTO
LECF_ALEAU
PAT2_CAEEL
COTN_BACSU
RECA_BORPE
                                                                                                                                                    HUG1_HUMAN
GUNA_MICBI
KV6K_MOUSE
                                                                                                                                                                                                            XYN1_ASPAW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; M16185; AAA22496.1; ALT_INIT
                                                                                                                                                                                                                                                                                                                                                                                                01-AUG-1988 (Rel. 08, Created)
01-NOV-1991 (Rel. 20, Last sequ
01-OCT-1994 (Rel. 30, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      55187 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 43.0
Matches 37; Conservative
                                                                                                                                                                                                                                                                                                                                                              STANDARD;
   120
338
342
1157
1157
261
261
352
352
362
108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bacillus subtilis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    [1]
SEQUENCE FROM N.A.
   13.00
12.77
12.77
12.77
12.77
12.66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=1423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BGLC OR GLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (Celĺulase)
                                                                                                                                                                                                                                                                                                                                                              GUN1_BACSU
P07983;
   CHAIN
ACT_SITE
ACT_SITE
DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                         RESULT 1
GUN1_BACSU
   4556666444444
45666666444444
                                                                                                                                                                                                                                                                                                                                                                P07983 bacillus su
P10475 bacillus su
P23549 bacillus su
P22533 caldocellum
                                                                                                                          2, 2002, 09:18:04; Search time 28.05 Seconds (without alignments) 121.473 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     erwinia car
erwinia car
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        clostridium clostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    m bifunctio
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     mus musculu
escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  caldocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  c endogluca
paenibacill
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         clostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             clostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               clostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                clostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           thermoplasm
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       rhizobium m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      paenibacill
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   erwinia car
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      methanococc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               schistosoma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     h peroxisom
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              rhodobacter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Description
                                                                                                                                                                                                      US-09-917-376-5
469
1 VSGGVKVQYKNNDSAPGDNQ......IRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            P22534
P10474
P204719
O02934
O02934
O06851
C59395
P5090
P5000
P5000
P5000
P5000
P5000
P5000
P5000
P5000
P500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q9upv9
Q15648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      P29352
P09815
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P14428
P03991
P32138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     P33693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            P50937
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P38108
                   GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                              hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                         105224 segs, 38719550 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GUX2_CLOSR
GUNZ_CLOSR
CBPA_CLOCL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FLA1_THEVO
TY3H_SCHMA
YA42_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HA10_MOUSE
HA1W_MOUSE
YIHQ_ECOLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GUNA_CALSA
GUNB_CALSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GUNI_CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CIPA_CLOTM
GUNN_ERWCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FLA1_METVA PUR9_MYCTU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EXOK_RHIME PIN8_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HV39_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PPRB_HUMAN

    protein search, using sw model

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    YCEA_PAELA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GUNA_PAELA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ERWCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GUNW_ERWCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATC5_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ICEN PSEFL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HIS6_RHOSH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MUCB_PSEAE
                                                                                                                                                                                                                                                                                                     , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GUNV
                                                                                                                                                                                                                                                                                                                                                                                                                    length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SwissProt_40:*
                                                                                                                                                                                                                                                                                 BLOSUM62
Gapop 10.0 ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1331
1742
1039
145
700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     July
                                                                                                                                                                                                                                                                                                                                                                              Potal number of
                                                                                                                                                                                                                                                                                                                                                                                                                      sed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 190
1186.5
1186.5
1186.5
1186.5
1186.5
1180.5
1190.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.5
1147.
                                                                                                                                                                                                                             score:
                                                                                                                                                                                                                                                                                   Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                  Minimum DB
Maximum DB
                                                                                             OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Database :
                                                                                                                                                                                                                                               Sequence:
                                                                                                                                                                                                                                                                                                                                         Searched:
                                                                                                                                                                                                          Title:
Perfect
                                                                                                                                  Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Š.
```

; ;

```
g
                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                               Lindahl V., Aa K., Tronsmo A.;
"Nucleotide sequence of an endo-beta-1,4-glucanase gene from Bacillus
subtilis CK-2.";
            62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            linkages in cellulose.
4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA
                                                                                                                                                                                                                                                                                           STRAIN-PAPILS;
MEDLINE-87066783; PubMed-3024130;
MacKay R.M., Lo A., Willick G., Zuker M., Baird S., Dove M.,
Moranelli F., Seligy V.;
Seligy V.;
"Structure of a Bacillus subtilis endo-beta-1,4-glucanase gene.";
Nucleic Acids Res. 14:9159-9170(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Aa K., Flengsrud R., Lindahl V., Tronsmo A.;
Characterization of production and enzyme properties of an
endo-beta-1,4-glucanase from Bacillus subtilis CK-2 isolated from
compost soil."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Antonie Van Leeuwenhoek 66:319-326(1994).
-1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
                                                                                                                                                  01-JUL-1989 (Rel. 11, Created)
01-JUL-1989 (Rel. 11, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Endoglucanase precursor (ES 3.2.1.4) (Endo-1,4-beta-glucanase)
(Carboxymethyl-cellulase) (CMCASE) (Cellulase).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Rose M., Entian K.;
Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                 Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
NCBL_TaxID=1423;
                                                                                                                               499 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                Antonie Van Leeuwenhoek 66:327-332(1994).
                                                           410 QMGCGNLTHKEVTLHKPKQGADTYLE 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; Z29076; CAA82317.1; -.
EMBL; X04689; CAA28392.1; -.
EMBL; X67044; CAA47429.1; -.
EMBL; Z73234; CAA97610.1; ALT_INIT.
EMBL; Z99113; CAB13696.1; ALT_INIT.
                                             63 AMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-CK-2;
MEDLINE-9522565; PubMed-7710280;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-95225655; PubMed-7710279;
                                                                                                                                STANDARD;
                                                                                                                                                                                                              BGLC OR GLD OR EGLS.
Bacillus subtilis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PIR; A26114; A26114.
HSSP; O85465; 1A3H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 30-45.
                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HYDROLASES)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN=168;
                                                                                                                                GUN2_BACSU
                                                                                                                                        P10475;
                                                                                                                  GUN2_BACSU
                    a
                                              ò
                                                                   요
                                                                                                                                ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       linkages in cellulose.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Park S.H., Kim H.K., Pack M.Y.;
"Characterization and structure of the cellulase gene of Bacillus subtilis BSE616.";
Agric. Biol. Chem. 55:441-448(1991).
-!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
                                                                                                                                                                                                                                               PROTON DONOR (BY SIMILARITY).
NUCLEOPHILE (BY SIMILARITY).
CELLULOSE-BINDING (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1991 (Rel. 20, Last sequence update)
01-NOV-1991 (Rel. 20, Last sequence update)
01-FEB-1995 (Rel. 31, Last annotation update)
Endoglucanase precursor (EC 3.2.1.4) (Endo-1.4-beta-glucanase)
(Carboxymethyl-cellulase) (CMCASE) (Cellulase).
                                                                                                                                                                                                                                                                                                                                                                                                              Length 499;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                        Probom; PD001947; CBD 3; 1.
PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
Cellulose degradation; Hydrolase; Glycosidase; Signal;
                                                                                                                                                                                                                                                                                                                  -> N (IN REF. 2).
8F735FF711B3EAE2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                         Bacteria, Firmicutes, Bacillus/Clostridium group, Bacillus/Staphylococcus group; Bacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   499 AA
                                                                                                                                                                                                                            ENDOGLUCANASE.
                InterPro; IPR001956; ČBD_3.
InterPro; IPR001547; Glyco_hydro_F5.
Pfam; PF00942; CBD_3; l.
Pfam; PF00150; cellulase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PIR; JN0111; JN0111.
HSSP; O85465; 1A3H.
InterPro; IPR001956; CBD_3.
InterPro; IPR001547; Glyco_hydro_F5.
Pfam; PF00942; CBD_3; 1.
Pfam; PF00150; cellulase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          410 QIGCGNVTHKFVTLHKPKQGADTYLE 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=91299280; PubMed=1368694;
                                                                                                                                                                                                                                                                                                                         ß
                                                                                                                                                                                                  29
499 ED
169 PF
257 NC
499 CF
283
55287 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; D01057; BAA00859.1; -.
Subtilist; BG10437; bglC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                          30
169
257
350
283
499 AA;
                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity
Matches 35; Conserv
                                                                                                                                                                                 Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bacillus subtilis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=1423;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HYDROLASES)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-BSE616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GUN3_BACSU
ID GUN3_BACSU
AC P23549;
                                                                                                                                                                                                                                                                         ACT_SITE
DOMAIN
                                                                                                                                                                                                                                                  ACT_SITE
                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                    CONFLICT
                                                                                                                                                                                                       SIGNAL
                                                                                                                                                                                                                               CHAIN
```

us-09-917-376-5.rsp

```
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                           ACT_SITE
                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                            SIGNAL
                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GUNA_CALSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                 ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY FRAMILY 5 OF GLICOSYL HYDROLASES).
SIMILARITY: IN THE C-PERMINAL SECTION; BELONGS TO CELLULASE FAMILY J (FAMILY 44 OF GLYCOSYL HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 terium 'Caldocellum saccharolyticum'.";
1. Environ. Microbiol. 57:694-700(1991).
PUNCTION: DEGRADATION OF HEMICELLULOSES, THE SECOND MOST ABUNDANT POLYSACCHARIDES IN NATURE. CONTAINS TWO CATALYTIC DOMAINS WITH MANNANASE AND ENDOGLUCANASE ACTIVITIES.
CATALYTIC ACTIVITY: Random hydrolysis of 1,4-beta-D-mannosidic
                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                      01-AUG-1991 (Rel. 19, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2001 (Rel. 40, Last annotation update)
beta-mannanase/endoglucanase A precursor [Includes: Mannan endo-1,4-beta-mannosidase A (EC 3.2.1.78) (Beta-mannanase) (Endo-1,4-mannanase); Endo-1,4-beta-glucanase (EC 3.2.1.4) (Cellulase)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -!- MISCELLANEOUS: THIS ENZYME IS MOST ACTIVE AT PH 6 AND 80 DEGREES CELSIUS.
                                                                                                                                                                                                                             4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              [2]

REQUENCE OF 1-346 FROM N.A.

MEDLINE-91247819; Pubmed-2039230;

Luethi E., Jasmat N.B., Grayling R.A., Love D.R., Bergquist P.L.;

"Cloning, sequence analysis, and expression in Escherichia coli of gene coding for a beta-mannanase from the extremely thermophilic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
Bacteria; Firmicutes; Bacillus/Clostridium group;
Thermoanaerobacter group; Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 part of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          linkages in mannans, galactomannans, glucomannans, and galactoglucomannans.
CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
                                                                       PROTON DONOR (BY SIMILARITY).

WOLGBOPHIE (BY SIMILARITY).

CELLULOSE-BINDING (BY SIMILARITY).

ZERZIEJDBBBACAO4 CRC64;
                                                                                                                                                             40.1%; Score 188; DB 1; Length 499; 40.0%; Pred. No. 7.4e-14; ive 18; Mismatches 31; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-93119139; PubMed-1476429;
Gibbs M.D., Saul D.J., Luthi E., Bergquist P.L.;
"The beta-mannanase from 'Caldocellum saccharolyticum' is
                                Cellulose degradation; Hydrolase; Glycosidase; Signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Appl. Environ. Microbiol. 58:3864-3867(1992)
[2]
                                                                                                                                                                                                                                                                                                                                                                                             PRT; 1331 AA
                                                           ENDOGLUCANASE
CBD_3; 1.
GLYCOSYL_HYDROL_F5; 1.
                                                                                                                                                                                                                                                                                                       411 LGCGNVTYKFVTLHKPKQGADTYLE 435
                                                                                                                                                                                                                                                                                     64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                    55169 MW;
                                                                                                                                                                              Best Local Similarity 40.0% Matches 34; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linkages in cellulose.
                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                           29
169
257
257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 multidomain enzyme.";
                                                                                     257
350
499 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=44001;
                                                                                                                                                                                                                                                                                                                                                                                             MANB_CALSA
                                                                       ACT_SITE
ACT_SITE
DOMAIN
                                                                                                                    SEQUENCE
                                                                                                                                                                   Query Match
                                             SIGNAL
                                                            CHAIN
 ò
                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                  g
```

```
1;
modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 1516-1742 FROM N.A.
MEDLINE-91247819; PubMed-2039230;
Leuthi E., Jasmat N.B., Grayling R.A., Love D.R., Bergquist P.L.;
"Cloning, sequence analysis, and expression in Escherichia coli of a gene coding for a beta-mannanse from the extremely thermophilic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
Bacteria; Firmicutes; Bacillus/Clostridium group;
Thermoanaerobacter group; Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-OCT-1996 (Rel. 34, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
Endoglucanase A precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase A)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRO/SER/THR-RICH (PT BOX).
SUBSTRATE-BINDING (POTENTIAL).
PRO/SER/THR-RICH (PT BOX).
CATALYTIC (ENDOGLUCANANE ACTIVITY).
BROYON DONOR (BY SIMILARITY).
NUCLEOPHILE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-95336703; PubMed-7612247;
Te'O V.S., Saul D.J., Bergquist P.L.;
"celA, another gene coding for a multidomain cellulase from the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               39.8%; Score 186.5; DB 1; Length 1331; 42.5%; Pred. No. 3.5e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BETA-MANNANASE/ENDOGLUCANASE A.
CATALYIC (MANNANASE ACTIVITY).
PRO/SER/THR-RICH (PT BOX).
SUBSTRATE-BINDING (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PTPT -> RQHQHRQ (IN REF. 2). FFBCA51BB8D8F0E0 CRC64;
                                                                                                                                                                                                                                                                                                                                           InterPro: IPR001956; CBD_3.
InterPro: IPR001547; GBD_3:
Pfam: PF00150; CBIJUlase; 1.
Pfam: PF00150; cellulase; 1.
Probom: PD001947; CBD_3; 2.
PROSTE: PS00659; GLYCGSYL_HYDROL_F5; 1.
Hydrolase; Glycosidase; Cellulose degradation; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   extreme thermophile Caldocellum saccharolyticum."; Appl. Microbiol. Biotechnol. 43:291-296(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               T -> P (IN REF. 2).
TPTPTPT -> ROHOHRO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1742 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               423 AQIGASNVTFKFVKLSSSVSGADYYLE 449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62 AAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-AUG-1991 (Rel. 19, Created)
01-OCT-1996 (Rel. 34, Last sea
                                                                                                                                           EMBL; L01257; AAA71887.1; -. EMBL; M36063; AAA72861.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   146892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Multifunctional enzyme
                                                                                                                                                                                                                               PIR; B43745; B43745.
PIR; A48954; A48954.
HSSP; Q06851; 1NBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=44001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           326
362
519
565
721
781
162
257
257
338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (Cellulase A)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GUNA_CALSA
P22534;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HAD DE POLITION OF THE PRINCE ```

Created)

```
(EC 3.2.1.4) (Endo-1,4-beta-glucanase) (Cellulase) (Cellobiohydrolase); Exoglucanase (EC 3.2.1.91) (Exocellobiohydrolase)
              01-JUL-1989 (Rel. 11, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Endoglucanase/exoglucanase B precursor (Includes: Endoglucanase
 (Rel.
  01-JUL-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BYBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ä
                                                                                                                       SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO CELLULASE FAMILY E (FAMILY 9 OF GLYCOSYL HYDROLASES).
SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY L (FAMILY 48 OF GLYCOSYL HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
          Appl. Environ. Microbiol. 57:694-700(1991).
-!- FUNCTION: THE N-TERMINAL DOMAIN OF CELA ENCODES FOR AN
ENDOGLUCARNSE ACTIVITY ON CARBOXWETHYLCELLULOSE. THE C-TERMINAL
DOMAIN PROBABLY ACT SYNERGISTICALLY TO HYDROLYZE CRYSTALLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                    linkages in cellulose.
PTM: THE LINKER REGION (ALSO TERMED "HINGE") MAY BE A POTENTIAL
SITE FOR PROTEOLYSIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                        CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LINKER ("HINGE") (PRO-THR BOX).
CELLULOSE-BINDING (BY SIMILARITY).
LILKUR ("HINGE") (PRO-THR BOX).
CELLULOSE-BINDING (BY SIMILARITY).
LINKER ("HINGE") (PRO-THR BOX).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 1; Length 1742;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Signal; Repeat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3F0699A2123EED07 CRC64;
                                                                                                                                                                                                                                                                                                                                             InterPro: IPR001356; CBD_3.
InterPro: IPR001356; Glyco_hydro_48.
InterPro: IPR0010556; Glyco_hydro_48.
InterPro: IPR00101; Glyco_hydro_9.
Pfam: PF00104; CBD_3; 3.
Pfam: PF02011; Glyco_hydro_48; 1.
Pfam: PF00159; Glyco_hydro_9; 1.
PRIME: PR00084; GLHVELASE48.
ProDom: PD011903; Glyco_hydro_48; 1.
PROSITE; PS00699; GLYCOSL_HYDROL_F9_1; 1.
PROSITE; PS00698; GLYCOSL_HYDROL_F9_2; 1.
CGllulose degradation; Hydrolase; Glycosidase; Sig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    39.8%; Score 186.5; DB 1;
42.5%; Pred. No. 4.9e-13;
tive 18; Mismatches 31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A (IN REF. 2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ENDOGLUCANASE A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CATALYTIC 2.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
bacterium 'Caldocellum saccharolyticum'.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CATALYTIC 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            763 AQIGASNVTFKFVKLSSSVSGADYYLE 789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                 EMBL; M36063; AAA72860.1; -
EMBL; LO1257; -; NOT_ANNOTATED_CDS.
PIR; A47745; A43745
HSSP; P26221; 1TF4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ٨
                                                                                                                                                                                                                                                                                    EMBL; L32742; AAA91086.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 193696
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Simi
Matches 37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Cellulose o
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
ACT_SITE
ACT_SITE
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic linkages in cellulose.
-!- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glucosidic linkages in cellulose and cellotetraose, releasing cellobiose from the non-reducing ends of the chains.
-!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY F (FAMILY 10 OF GLYCOSYL HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDIATE-89081898; PubMed-2789517;
MEDIATE-89081898; PubMed-2789517;
Saul D.J., Williams L.C., Love D.R., Chamley I.W., Bergquist P.I.;
Nucleotide sequence of a gene from Caldocellum saccharolyticum
encoding for exocellulase and endocellulase activity.";
Nucleic Acids Res. 17:439-439(1989).

-i- FONCTION: THIS PROFEIN IS MADE UP OF TWO DOMAINS: THE N-TERMINAL
DOMAIN HAS EXOGLUCANASE ACTIVITY WHILE THE C-TERMINAL DOMAIN IS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA 62
                                                                                                  Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
Bacteria; Firmicutes; Bacillus/Clostridium group;
Thermoanaerobacter group; Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          THR/PRO-RICH, TANDEM REPEATS OF T-P. CELLULOSE-BINDING (BY SIMILARITY). THR/PRO-RICH, TANDEM REPEATS OF T-P. PROTON DONOR (POTENTIAL).
NUCLEOPHILE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 1039;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ENDOGLUCANASE/EXOGLUCANASE B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HSSP, OG6851, DBC.
InterPro: IPR001956; CBD_3.
InterPro: IPR001956; CBD_3.
InterPro: IPR001056; CBD_3.
InterPro: IPR001054; Glyco_hydro_10.
InterPro: IPR001547; Glyco_hydro_F5.
InterPro: IPR00134; Glyco_hydro_10.
InterPro: IPR003134; Glyco_hydro_10.
IPR00315; PR00134; Glyco_hydro_10.
IPR00315; PR001947; CBD_3; 1.
IPR05ITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
IPR0SITE;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY.
0E0378171594DDAE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    38.9%; Score 182.5; DB 1
41.9%; Pred. No. 7.5e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Signal
(1,4-beta-cellobiohydrolase)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; X13602; CAA31936.1; -. PIR; S02711; S02711.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      117641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Multifunctional enzyme;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1039
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AN ENDOGLUCANASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              177 17
285 28
792 79
1039 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 36; Conserv
                                                                                                                                                                                                                                                               NCBI_TaxID=44001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

ï

1039 AA.

STANDARD;

GUNB_CALSA P10474;

A D

GUNB_CALSA

RESULT 7
YCEA_PAELA

å g DDTTDDDTTDDDTTDDTTDTTMNX

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics. The There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
01-NOV-1995 (Rel. 32, Last annotation update)
Endoglucanase I precursor (EC 3.2.1.4) (EGI) (Endo-1,4-beta-glucanase)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                     1 VSGGVKVQYKNND-SAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNC 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "Gene sequence and properties of Cell, a family E endoglucanase from
                                                                                                                              Hansen C.K., Diderichsen B., Joergensen P.L.; "celA from Bacillus lautus PL236 encodes a novel cellulose-binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Clostridium thermocellum.
Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
Clostridium.
                                                                                                                                                                                 endo-beta 1,4 glucanase.";
J. Bacteriol. 174:3522-3531(1992).
-!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 700;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hazlewood G.P., Davidson K., Laurie J.I., Huskisson N.S., Gilbert H.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Přám; Př00942; CBD_3; 1.
ProDom; PD001947; CBD_3; 1.
Cellulose degradation; Hydrolase; Glycosidase; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BY SIMILARITY.
3D5C8CADA53EEE0F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      36.7%; Score 172; DB 1;
44.9%; Pred. No. 7.4e-12;
tive 18; Mismatches 27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ENDOGLUCANASE A.
CELLULOSE-BINDING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           879 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A., AND SEQUENCE OF 56-69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ||| :| ||: ||| :|| :|| :
DWAKLGGSNIQISFGNHNGA--DSDTYAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             60 DWAAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                       MEDLINE-92276330; PubMed-1592807;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-NCIB 10682;
MEDLINE=93171873; PubMed=8436949;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        76910 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; M76588; AAA22303.1; -.
PIR; B41897; B41897.
PIR; S27499; S27499.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HSSP; Q06851; 1NBC.
InterPro; IPR001956; CBD_3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           33
700
700
213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  548
213
700 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=1515;
NCBI_TaxID=1401;
                                                                                                                                                                                                                                                                                                                                                                                     HYDROLASES)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query ...
Best Local Similar
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (Cellulase I)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GUNI_CLOTM
002934;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GUNI_CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
NAME OF THE PROPERTY OF THE PR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lab-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hansen C.K., Joergensen P.L., Diderichsen B.;
"celA from Bacillus lautus PL236 encodes a novel cellulose-binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
Endoglucanase A precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 1; Length 145;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             145 AA; 15782 MW; 9514E3A71B106AEB CRC64;
                                                                                                                                                                                                                                                               01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
16-027-2001 (Rel. 40, Last annotation update)
Hypothetical protein in CELA 5'region (Fragment).
Paentbacillus lautus (Bacillus Jautus).
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Paenibacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Paenibacillus lautus (Bacillus lautus).
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Paenibacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 180.5; DB 1;
Pred. No. 1.2e-13;
3; Mismatches 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            700 AA.
                                                                                                                                                                                                                 145 AA.
                                                      :| |: : : || |: OIGASNVTFKFVKLSSSVSGADYYLE 504
                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          endo-beta-1,4-glucanase.";
J. Bacteriol. 174:3522-3531(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN=PL236;
MEDLINE=92276330; PubMed=1592807;
                    63 AMGCGNIRASFGSVNPATPTADTYLO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SNIQISFG--NHTGTNSDTYVE 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          38.5%;
|larity 43.9%;
|Conservative 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; M76588; AAA22302.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  67 GNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HSSP, Q06851; INBC.
InterPro; IPR001956; CBD_3.
Pfam; PF00942; CBD_3; 1.
ProDom; PD001947; CBD_3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (Cellulase A) (EG-A).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hypothetical protein.
NON_TER 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PIR; A41897; A41897.
PIR; S27498; S27498.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=1401;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Simi
Matches 36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GUNA_PAELA
P29719;
                                                                                                                                                                                                                 YCEA_PAELA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
```

RESULT 8 GUNA_PAELA

9

g ŏ g

ò

3,

9

```
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NON_TER
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REPEAT
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           006851;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CIPA_CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     : | | :|| | : | | | : | | | : | | | 36 IKGEVVLQYANGNAGATSNSINPRFKIINNGTKAINLSDVKIRYYYTKEGGASONFW-CD 794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-0CT-1994 (Rel. 30, Created)
01-0CT-1994 (Rel. 30, Last sequence update)
16-0CT-2001 (Rel. 40, Last annotation update)
Cellulosomal scaffolding protein B (Cellulosomal glycoprotein S1/SL)
(Cellulose integrating protein B) (Fragment).
Clostridium thermocellum.";
J. Gen. Microbiol. 139:307-316(1993).
-!- FUNCTION: THE ENZYME CATALYZES THE ENDOHYDROLYSIS OF 1,4-BETA-GLUCOSIDIC LINRAGES IN CELLULOSE, LICHEMIN AND CERRAL BETA-GLUCOSIDIC LINRAGES IN CELLULOSE, LICHEMIN AND CERRAL BETA-D-GLUCANS. PRINCIPALLY ACTIVE AGAINST BARLEY BETA-GLUCAN.
                                                                                                                                                                              GLYCOSYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
Clostridium.
                                                                                                         CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        35.3%; Score 165.5; DB 1; Length 879;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-93146373; PubMed=1490597;
Poole D.M., Morag E., Lamed R., Bayer E.A., Hazlewood G.P.,
Gilbert H.J.;
                                                                                                                                                                          SIMILARITY: BELONGS TO CELLULASE FAMILY E (FAMILY 9 OF HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. 5.3e-11;
; Mismatches 32; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CELLULOSE-BINDING (BY SIMIL BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
3. 35A600694514A927 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                         HSSP, P26221; TF4.
InterPro; IPR001956; CBD_3.
InterPro; IPR001956; CBD_3.
Pfam; PF00942; CBD_3; 2.
Pfam; PF000759; G1YCo_hydro_9; 1.
ProDom; PD001947; CBD_3; 1.
PROSITE; PS00592; GLYCOSVL_HYDROL_F9_1; 1.
PROSITE; PS00698; GLYCOSYL_HYDROL_F9_2; 1.
Cellulose degradation; Hydrolase; Glycosidase; SigsIGNL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ENDOGLUCANASE I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   772 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CATALYTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                    linkages in cellulose.
PATHWAY: CELLULOSE DEGRADATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ..
Σ
                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; L04735; AAA20892.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              33.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           96116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Clostridium thermocellum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR; A47704; A47704.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      729
448
486
495
879 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=1515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CIPB_CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local S:
Matches 29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
DOMAIN
ACT_SITE
ACT_SITE
ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        001866;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 10
CIPB_CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
    SOUTH THE THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
between the SWISS Institute of Bioinformatics and the EMBL outstation—the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration
"Identification of the cellulose-binding domain of the cellulosome subunit S1 from Clostridium thermocellum YS.";
FEMS Microbiol. Lett. 78:181-186(1992).

-!- FUNCTION: ACTS AS A SCAPFOLDING PROTEIN IN THE CELLULOSOME. IT PROMOTES BINDING OF CELLULOSE TO THE CATALYTIC DOMAINS OF THE CELLULOLYTIC ENZYMES PROBABLY THROUGH THE BINDING OF THE NIBE REPEATED DOMAINS WITH THE DUPLICATED SEGMENT OF 24 AMINO ACIDS PRESENT IN CATALYTIC SUBUNITS OF THE CELLULOSOME.
-!- SUBCELLULAR LOCATION: CELL SURFACE.
-!- SIMILARITY: A 24 RESIDUES DOMAIN IS REPEATED TWICE IN THIS ENZYME.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            35.2%; Score 165; DB 1; Length 772;
36.1%; Pred. No. 5.2e-11;
Live 23; Mismatches 29; Indels 10; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Clostridium thermocellum.
Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
Clostridium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JUN-1994 (Rel. 29, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
01-MAR-2002 (Rel. 41, Last annotation update)
cellulosomal scaffolding protein A precursor (Cellulosomal glycoprotein S1/SL) (Cellulose integrating protein A) (Cohesin).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 X 147 AA APPROXIMATE REPEATS.
1 (PARTIAL).
2.
3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRO/THR-RICH.
2 X 24 AA APPROXIMATE REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF00942; CBD_3; 1.
Pfam; PF00963; Cohesin; 3.
Pfam; PF00464; Dockerin_1; 2.
Prodom; PD001947; CBD_3; 1.
PROSITE; PS00448; CLOS_CELLULOSOME_RPT; 2.
Cellulose degradation; Cell wall; Glycoprotein; Repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BBF06DE5E094FE10 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CELLULOSE-BINDING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRO/THR-RICH.
PRO/THR-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro: IPR001956; CBD_3.
InterPro: IPR002102; Cohesin.
InterPro: IPR002105; Dockerin_1.
InterPro: IPR002048; EF-hand.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        82491 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; X68233; CAA48312.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1

<1

<1

<1

462

462

81

273

440

710

710

710

743

772 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity
Matches 35; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HSSP; 006851;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CIPA_CLOTM
```

; ;

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δλ
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                 SEQUENCE OF 1820-1853 FROM N.A.
MEDILINE-9230931; PubMed=8458832;
Fujino T., Beguin P., Aubert J.-P.,
"Organization of a clostridium thermocellum gene cluster encoding the cellulosomal scaffolding protein CipA and a protein possibly involved in attachment of the cellulosome to the cell surface.";
J. Bacteriol. 175:1891-1899(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELLULOLYTIC ENZYMES.
SUBCELLULAR LOCATION: CELL SURFACE.
DOMAIN: CONTAINS 9 COPIES OF A DOMAIN (COHESIN) OF ABOUT 150
RESIDUES THAT BINDS TO THE DOCKERIN DOMAIN BORN BY THE CATALYTIC
COMPONENTS OF THE CELLULOSOME.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN: A 24 RESIDUES DOMAIN IS REPEATED TWICE IN THIS ENZYME AS WELL AS IN OTHER C.THERMOCELLUM CELLULOSOME ENZYMES. THIS DOMAIN MAY FUNCTION AS THE BINDING LIGAND FOR THE SL COMPONENT.
                                                                                                                                                                                                                                                                                                                                                                        Shimon L.J.W., Bayer E.A., Morag E., Lamed R., Yaron S., Shoham Y.,
                                                                                                                 "Sequencing of a Clostridium thermocellum gene (cipA) encoding the cellulosomal SL-protein reveals an unusual degree of internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               J. Mol. Biol. 273:701-713(1997).
-!- FUNCTION: ACTS AS A SCAFFOLDING PROTEIN IN THE CELLULOSOME. IT PROMOTES BINDING OF CELLULOSE TO THE CATALYTIC DOMAINS OF THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 368-522.
MEDLINE-97076134; Pubmed-8918451;
Tormo J., Lamed R., Chirino A.J., Morag E., Bayer E.A., Shoham
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               'Crystal structure of a bacterial family-III cellulose-binding
                                                                               Gerngross U.T., Romaniec M.P.M., Kobayashi T., Huskisson N.S.,
                                                                                                                                                                                                                                                                                                                                                                                                     A cohesin domain from Clostridium thermocellum: the crystal
structure provides new insights into cellulosome assembly.";
Structure 5:381-390(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE=98022914; PubMed=9402065;
Tavares G.A., Beguin P., Alzari P.M.;
"The crystal structure of a type I cohesin domain at 1.7-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               domain: a general mechanism for attachment to cellulose.";
EMBO J. 15:5739-5751(1996).
                                 SEQUENCE FROM N.A., AND SEQUENCE OF 29-40 AND 42-43,
                                                                                                                                                                                                                                                                                                                                       X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 184-321.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 1213-1361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; L08665; -; NOT_ANNOTATED_CDS.
EMBL; X67506; CAA47840.1; -.
PIR; S36859; S36859.
                                                 STRAIN-ATCC 27405;
MEDLINE-93302508; PubMed-8316083;
                                                                                                                                                                                                                                                                                                                                                          MEDLINE=97238934; PubMed=9083107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002105; Dockerin_1.
InterPro; IPR002048; EF-hand.
                                                                                                                                                                     Mol. Microbiol. 8:325-334(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001956; CBD_3.
InterPro; IPR002102; Cohesin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1AOH; 08-JUL-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      26-SEP-97.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1ANU; 23-JUL-97
                                                                                                   Demain A.L.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     resolution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1NBC;
                                                                                                                                                     homology.
                                                                                                                                                                                                                                                                                                                                                                                            Frolow F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Steitz
```

Pfam; PF00942; CBD_3; 1. Pfam; PF00963; Cohesin; 9.

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the PMBL outstation between the informatics Institution. There are no restrictions on its use by non-profit institutions and a st its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                    5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN=ATROSEPTICA FCBR C18;
MEDLINE-98299944; PubMed-9636315;
Olsen O., Thomsen K.K., Weber J., Duus J.O., Svendsen I., Wegener C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
                                                                                                        CELLULOSOMAL SCAFFOLDING PROTEIN A. CELLULOSE-BINDING (BY SIMILARITY). 9 X 150 AA APPROXIMATE TANDEM REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "Transplanting two unique beta-glucanase catalytic activities into one multienzyme, which forms glucose.";
Biotechnology 14:71-76(1996).
-:- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 VSGCVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-NoV-1997 (Rel. 35, Created)
01-NoV-1997 (Rel. 35, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
Endoglucanase N precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Pfam; PF00404; Dockerin_1; 2.
ProDom; PD001947; CBD_3; 1.
PROSITE; PS00018; EF_HAND; UNKNOWN_1.
PROSITE; PS00448; CLOS_CELLULOSOME_RPP; 2.
Cellulose degradation; Cell Wall; Glycoprotein; Repeat; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                    10;
                                                                                                                                                                                                                                                                                                                                                                                                     35.2%; Score 165; DB 1; Length 1853; 36.1%; Pred. No. 1.5e-10; cive 23; Mismatches 29; Indels 10
                                                                                                                                                                                                                                                                                              24 AA APPROXIMATE REPEATS
                                                                                                                                                                                                                                                                                                                                         -> AA (IN REF. 1).
3ABDDC03ABFC5372 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         444 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 WAAM-----GCGNIRASFGSVNPATPTADTYLO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linkages in cellulose.
SUBCELLULAR LOCATION: Secreted.
                                                                                                                                        9 x

1-1.

1-2.

1-4.

1-5.

1-7.

1-8.
                                                                                                                                                                                                                                                                                                                                             MW;
                                                                                                                                                                                                                                                                                                                                                           196831
                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                  1031
1196
1361
1526
1690
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Erwinia carotovora.
                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        von Wettstein D.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HYDROLASES).
                                                                                                                                                                                                                                                               1384
1548
1791
1791
1824
1615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pectobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=554;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (Cellulase N)
                                                                              3D-structure
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GUNN_ERWCA
Q59394;
                                                                                                                                                                                                                                                                                                                                         CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                    REPEAT
REPEAT
                                                                                                                                                                                                                                                               REPEAT
REPEAT
                                                                                              SIGNAL
                                                                                                                           DOMAIN
                                                                                                                                          DOMAIN
                                                                                                                                                            REPEAT
                                                                                                                                                                        REPEAT
                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                       REPEAT
                                                                                                                                                                                                                                                   REPEAT
                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                REPEAT
                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
```

```
g
                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                               οy
entilles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                               'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MOI. Gen. Genet. 241:341-350(1993).
-!- FUNCTION: ENDOGLUCANASE WITH SOME EXOGLUCANASE ACTIVITY. THE PH
OPTIMUM IS ABOUT 7.0 AND THE TEMPERATURE OPTIMUM ABOUT 42 DEGREES
                                                                                                                                                                                                                                                                                                                                                                            357 TGDVVLQYRNVDNNPSDDAIRMAVNIKNTGSTPIKLSDLQVRYYFHDDGKPGANLFV--- 413
                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                            2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDG--GSSTLVYNC 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linkages in cellulose.
SUBCELLULAR LOCATION: Secreted.
SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           carotovora: evidence for an evolutionary 'mix-and-match' of enzyme
                              R REME; L139788; ALL.

R REME; L139788; ALL.

R HSSP, 085465; 1A3H.

DR InterPro; IPR001956; CBD_3.

DR Pfam; PF00942; CBD_3; 1.

DR ProDom; P0001947; CBD_3; 1.

DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.

DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.

ET SIGNAL

FT CHAIN 32 444 ENDOGLUCANNSE N.

FT CHAIN 32 444 ENDOGLUCANNSE N.

STTE 168 168 PROTON DONON (BY SIMILARITY).

ANDLEOPHILE (BY SIMILARITY).

ANDLEOPHILE (BY SIMILARITY).

ANDLEOPHILE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bacteria: Proteobacteria; gamma subdivision; Enteropacteriaceae; Pectobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
Endoglucanase V precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase (Cellulase V).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
                                                                                                                                                                                                                                                                                                                               ų,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Cooper V.J.C., Salmond G.P.C.; "Molecular analysis of the major cellulase (CelV) of Erwinia
                                                                                                                                                                                                                                                                                               31.9%; Score 149.5; DB 1; Length 38.6%; Pred. No. 1.6e-09; ive 19; Mismatches 30; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              505 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                          414 DWANVGPNNIVTSTGTPAASTDKANRYV 441
                                                                                                                                                                                                                                                                                                                                                                                                                       DWAAMGCGNIRASFGSVNPATPTADTYL 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=94067016; PubMed-8246888;
                                                                                                                                                                                                                                                                                                                             34; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Erwinia carotovora
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=554;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-SCRI193;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GUNV_ERWCA
Q47096;
                                                                                                                                                                                                                                                                                               Query Match
Best Local S
Matches 34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          domains."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GUNV_ERWCA
                                                                                                                                                                                                                                                                                                                                                                                                                     9
    ò
                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                     á
                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

EMBL; X76000; CAA53592.1; -.

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                         5
                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                         2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDG--GSSTLVYNC 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
Endoglucanase VI precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase VI)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mae A., Heikinheimo R., Palva E.T.; "Structure and regulation of the Erwinia carotovora subspecies carotovora SCC3193 cellulase gene celV1 and the role of cellulase in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 phytopathogenicity.";
Mol. Gen. Genet. 247:17-26(1995).
-!- CAPALYTUTY: Endohydrolysis of 1,4-beta-D-glucosidic
- Inkages in cellulose.
-!- SUBCELLULAR LOCATION: Secreted.
-!- SUBLEARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                     CELLULOSE-BINDING (BY SIMILARITY).
PROTON DONG (BY SIMILARITY).
NUCLEOPHILE (BY SIMILARITY).
DBEA9337BB4D2623 CRC64;
                                                                                                                                                                                                                                                                                                                         5,
                                                                                                                                                                                                                                                                                      Length 505;
                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                      Signal.
                                                                                                                                                                                                                                                                                        31.9%; Score 149.5; DB 1; 38.6%; Pred. No. 1.8e-09;
                                                                                                                                                                                                                                                                                                                         30;
                                                    Pfam; PFUULDU; CELLLIS, 1.
Probom; PD001947; CBD_3, 1.
PROSTIE; PS00659; GLXCOSYLHYDROL_F5; 1.
Cellulose degradation; Hydrolase; Glycosidase; 31
                                                                                                                                                                                                                                                                                                       ; Pred No. 1.8e-09; 19; Mismatches 30
                                                                                                                                    ENDOGLUCANASE V.
                                                                                                                                                     CATALYTIC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                          411 DWANVGPNNIVTSTGTPAASTDKANRYV 438
                                                                                                                                                                                                                                                                                                                                                                                                                          DWAAMGCGNIRASFGSVNPATPTADTYL 87
InterPro; IPR001956; CBD_3.
InterPro; IPR001547; Glyco_hydro_F5.
Pfan; PF00942; CBD_3; 1.
Pfan; PF00150; cellulase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR001956; CBD_3.
InterPro; IPR001547; Glyco_hydro_F5.
Pfam; PF00942; CBD_3; 1.
                                                                                                                                                                         LINKER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE=95231512; PubMed=7715600;
                                                                                                                                                                                                                                        54900 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; X79241; CAA55823.1; -. HSSP; O85465; 1A3H.
                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                       31
334
352
352
505
256
                                                                                                                                                   32
335
353
168
256
505 AA;
                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
Matches 34; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Erwinia carotovora,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (Cellulase V1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=554;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-SCC3193;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pectobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GUNW_ERWCA
Q59395;
                                                                                                                                                                                                   ACT_SITE
ACT_SITE
SEQUENCE
                                                                                                                                     CHAIN
DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GUNW_ERWCA
                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
```

InterPro; IPR000556; Glyco_hydro_48

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                     ;
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHARACTERIZATION.
STRAIN-NCIB 11754;
MEDLINE-91364686; PubMed-1909625;
Bronnenmeier K., Ruecknagel K.P., Staudenbauer W.L.;
Bronnenmeier E., Ruecknagel K.P., Staudenbauer W.L.;
Purification and properties of a novel type of
exo.1, 4beta-glucanase (avicelase II) from the cellulolytic
thermophile Clostridium stercorarium.";
Eur. J. Biochem. 200:379-385(1991).
-!-CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glucosidic linkages
in cellulose and cellotetraose, releasing cellobiose from the non-
reducing ends of the chains.
-!-SMHIARITY: BELONGS TO CELLULASE FAMILY L (FAMILY 48 OF GLYCOSYL
                                                                                                                                                                                                                                                                                                                                                                                                                    354 TGDIVLQYRNVDNNPSDDAIRMAFNIKNTGSTPIKLSDLQVRYYFHDDGKPGANLFV--- 410
                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                              2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDG--GSSTLVYNC 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
01-0V-1997 (Rel. 35, Last annotation update)
Exoglucanase II precursor (EC 3.2.1.91) (Exocellobiohydrolase II)
(1,4-beta-cellobiohydrolase II) (Avicelase II).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Clostridium stercorarium.
Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
                                                                                                                                                                            CELLULOSE-BINDING (BY SIMILARITY).
PROTON DONOR (BY SIMILARITY).
NUCLEDPHILE (BY SIMILARITY).
OD7ECF74781565FA CRC64;
                                                                                                                                                                                                                                                                                                                                                     5;
                                                                                                                                                                                                                                                                                                         Length 504;

        Pfam;
        PF00150;
        cellulase;
        1.

        ProDom;
        PD001947;
        CBD_3;
        1.

        PROSITE;
        PS00659;
        GLYCOSYL_HYDROL_F5;
        1.

        Cellulose
        degradation;
        Hydrolase;
        Glycosidase;

        SIGNAL
        31
        SQ
        POTENTIAL.

        CHAIN
        32
        54
        BUDOGGLUCANASE VI.

        DOMAIN
        33
        334
        CATALYIC.

        DOMAIN
        335
        352
        LINKER

        DOMAIN
        353
        504
        CELLIOSE-BINDING (BY SIMILARITY)

        ACT_SITE
        168
        168
        PROTON DONOR (BY SIMILARITY)

                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Staudenbauer W.L.; submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bronnenmeier K., Kundt K., Riedel K., Schwarz W.H.,
                                                                                                                                                                                                                                                                                                    Query Match 31.4%; Score 147.5; DB 1; Best Local Similarity 37.5%; Pred. No. 3.1e-09; Matches 33; Conservative 19; Mismatches 31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    914 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           60 DWAAMGCGNIRASFGSVNPATPTADTYL 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                          54963 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; Z69359; CAA93280.1; -.
HSSP; Q06851; 1NBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                   335
353
168
256
504 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A. STRAIN-NCIB 11754;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=1510;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HYDROLASES)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GUX2_CLOSR
                                                                                                                                                                                             ACT_SITE
ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NA TETETE
                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

InterPro; IPR001956; CBD 3

```
Gaps
                                                                                                                                                                                                                                                         1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                         ij
                                                                                                                                                                                       31.4%; Score 147.5; DB 1; Length 914; 35.2%; Pred. No. 6.2e-09; tive 21; Mismatches 35; Indels 1
                                                                                                                          34 914 EXOGLUCANASE II.
914 AA; 103020 MW; D0DB6017D6DFF82C CRC64;
           pfam; pF00942; CBD_3; 1. Pfam; pF00942; CBD_3; 1. Pfam; pF02011; G1yco_hydro_48; 1. PrNINTS; PR00844; G1HYDRLASE48. ProDom; PD001947; CBD_3; 1. ProDom; PD011903; G1yco_hydro_48; 1. Cellulose degradation; Hydrolase; G1ycosidase; Signal.signal.
                                                                                                                                                                                                                                                                                                                                           61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                         Search completed: July 2, 2002, 09:18:05 Job time: 373 sec
                                                                                                                                                                                                        Best Local Similarity 35.2
Matches 31; Conservative
                                                                                                                          CHAIN
SEQUENCE
                                                                                                                                                                                          Query Match
Ω
                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                            ò
```

1;

```
083012 bacillus sp
045430 bacillus sp
0917b10 acetivibrio
0917f5 streptomyce
0913j8 clostridium
09f6j9 bacteroides
045996 clostridium
045932 clostridium
04532 clostridium
04532 clostridium
046392 clostridium
04119 bacillus sp
091119 bacillus sp
09113 bacillus sp
09113 acillus sp
09132 homo sapien
09132 homo sapien
09133 hilham rat
099375 candida tro
099377 candida tro
099775 candida tro
099775 homolas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Alicyclobacillus group; Caldibacillus.
NCBI_TaxID=74586;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.

MEDLINE=20120520; PubMed=10653733;

MEDLINE=20120520; PubMed=10653733;

Sanna A., Gibbs M.D., Chin C.W.J., Nelson P.J., Bergquist P.L.;

"A gene encoding a novel multidomain beta-1,4-mannanase from
Caldibacillus cellulovorans and action of the recombinant enzyme (
kraft pulp.";

Microbiol. 66:664-670(2000).

EMBL; AFI63837; AAF222733.1; -.

HSSP; O06851; INBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 48.1%; Score 225.5; DB 2; Length 170; Best Local Similarity 47.7%; Pred. No. 2.6e-17; Matches 41; Conservative 17; Mismatches 27; Indels 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  170 AA; 18493 MW; 7AC9D33F44E3A0B4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
HYPOTHETICAL 18.5 KDA PROTEIN (FRAGMENT).
Caldibacillus cellulovorans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ALIGNMENTS
083012
095873
095875
095378
09773
09774
04599
09232
09332
09119
09119
09119
091132
091133
091133
091133
091133
091133
091133
091133
091133
091133
091133
091133
091133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR001956; CBD_3.
Pfam; PF00942; CBD_3; 1.
Probom; PD001947; CBD_3; 1.
Hypothetical protein.
NON_TER 1 1
SEQUENCE 170 AA; 18493 MW;
    55
112
113
113
114
115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
  501
4486
473
473
8386
11546
11546
11546
11546
11546
11546
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11746
11
\begin{smallmatrix} 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.4 \\ 7.
                                                                                                                                                                                            69.5
69
68.5
186
180
172.5
168
165.5
150.5
130.5
130.5
126.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q9RFX6
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                              2, 2002, 09:17:34 ; Search time 91.42 Seconds (without alignments) 166.523 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q9rfx6 caldibacill
Q9rfx5 caldibacill
Q91818 caldibacill
Q92411 bacillus sp
Q9aqq4 caldicellul
O24820 thermophili
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    caldicellul
caldicellul
caldicellul
anaerocellu
caldicellul
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        bacillus su
bacillus su
bacillus sp
bacillus su
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Description
                                                                                                                                                                                                                                                  1 VSGGVKVQYKNNDSAPGDNQ.....IRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09rfx5
091818
091818
094411
094820
094395
094397
094396
094396
093116
095314
093116
093116
093116
093116
                      GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                562222 segs, 172994929 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Listing first 45 summaries
                                                                                              OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                            Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q9REX5
Q9L8L8
Q9AQG4
Q9AQG4
Q9X3P5
Q9AQG7
Q9X3P6
P9G311
Q9X3P6
Q9AQH0
Q9AJTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q9RFX6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_mhc:*
sp_organelle:*
sp_phage:*
sp_plant:*
sp_rodent:*
sp_virus:*
sp_vertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sp_human:*
sp_invertebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_rvirus:*
sp_bacteriap:*
sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Post-processing: Minimum Match 0% Maximum Match 100%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_archea:*
sp_bacteria:*
sp_fungi:*
                                                                                                                                                                                                                                                                                                                                                                                                                            Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                              US-09-917-376-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SPTREMBL_19:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          921
997
11751
11000
11770
1261
11711
996
499
499
499
                                                                                                                                                                                                                                                                                          BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      448.1
446.2
445.7
441.0
441.0
440.8
440.7
440.5
39.9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    225.5
216.5
2214.5
204.5
204.5
192.5
191.5
191.5
190.5
190.5
190.5
190.5
190.5
190.5
190.5
190.5
190.5
190.5
190.5
190.5
190.5
                                                                                                                                                                                                                                    Perfect score:
                                                                                                                                                                                                                                                                                          Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ••
                                                                                                                                                                                                                                                      Sequence:
                                                                                                                                                                                                                                                                                                                                                Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Database
                                                                                                                                    Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Result
```

1;

2

g

```
Pastor F.I.J., Pujol X., Blanco A., Vidal T., Torres A.L., Dlaz P.; "Molecular cloning and characterization of a multidomain endoglucanase from Paenibacillus sp BP-23: evaluation of its performance in pulp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "A novel thermostable multidomain 1,4-beta-xylanase from 'Caldibacillus cellulovorans' and effect of its xylan-binding domain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Lappi. Microbiol. Biotechnol. 55:61-68(2001).

- SIMILARITY: CONNAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.

R BHBL, AJ133614, CAB38941.1; -.

R HSSP: P26221; 1TF4.

R InterPro; IPR001956; CBD_3.

R InterPro; IPR001956; CBD_3.

R InterPro; IPR0013961; FN_III.

R InterPro; IPR0013961; Enjoch, Jujoch, Ju
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   45.7%; Score 214.5; DB 2; Length 921;
46.5%; Pred. No. 3.3e-15;
tive 16; Mismatches 29; Indels 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL.
BETA-1,4-XYLANASE XYNA.
W; C5DDD1A7F7567413 CRC64;
                                                                                                                                                                                                History (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (1985) (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bacteria: Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              997 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PROSITE; PS00592; GLYCOSYL_HYDROL_F9_1; 1. PROSITE; PS00698; GLYCOSYL_HYDROL_F9_2; 1. PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-MAY-1999 (TrEMBLrel. 10, Created)
01-MAY-1999 (TrEMBLrel. 10, Last seq
01-DEC-2001 (TrEMBLrel. 19, Last ann
CELLULASE PRECURSOR (EC 3.2.1.4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        :|| |:| | : | | | 830 WVGCSNLRGSLVKLTTGRTGADYYLE 855
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN=BP-23;
MEDLINE=21129642; Pubmed=11234960;
                                                                                on enzyme activity.";
Microbiology 146:2947-2955(2000).
EMBL; AF200304; AAF61649.1; -.
HSSP; Q06851; 1NBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Glycosidase; Hydrolase; Signal, SIGNAL 1 34 POT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  34 921 BE
921 AA; 102380 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 46.5% Matches 40; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bacillus sp. BP-23.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      refining."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q924I1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0924I1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CELB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
    RAT RAT DE RAT D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Caldibacillus cellulovorans.
Bacteria: Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Alicyclobacillus group; Caldibacillus.
NCBI_TaxID=74586;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   XYNA.
Caldibacillus cellulovorans.
Bacteria: Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Alicyclobacillus group; Caldibacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
MEDLINE-20120520; Pubwed-10653733;
Sunna A., Gibbs M.D., Chin C.W.J., Nelson P.J., Bergquist P.L.,
*A gene encoding a novel multidomain beta-1,4-mannanase from
Caldibacillus cellulovorans and action of the recombinant enzyme on
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 46.2%; Score 216.5; DB 2; Length 930; Similarity 46.5%; Pred. No. 2e-15; 0; Conservative 17; Mismatches 28; Indels 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       930 MULTIDOMAIN BETA-1,4-MANNANASE.
101576 MW; 0086638D54D1A2CC CRC64;
                                                                                                                                                                                                                                                                                                                                 01-WAY-2000 (TrEMBLrel. 13, Created)
01-WAY-2000 (TrEMBLrel. 13, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
MULTIDOMAIN BETP-1,4-WANNANASE PRECURSOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001956; CBD_3.
InterPro; IPR001302; Chitin_bind_3.
InterPro; IPR001302; Chitin_bind_3.
InterPro; IPR0015047; Glyco_bydro_F5.
InterPro; IPR002965; P_rich_extensn.
Pfam; PF00142; CBD_3; 2.
Pfam; PF00150; cellulase; 1.
Pfam; PF00150; Chitin_bind_3; 1.
PRIMTS; PR01217; PRICHEXTENSN.
Probom; PD001947; CBD_3; 2.
PROSITE; PS00659; GLXCOSYL_HYDROL_F5; UNKNOWN_1.
                                                                                                                                                                                                                                                          930 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   921 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               kraft pulp.";
Appl. Environ. Microbiol. 66:664-670(2000).
EMBL; AF163837; AAF22274.1; -.
HSSP; Q06851; INBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDLINE-20519260; PubMed-11065373;
Sunna A., Gibbs M.D., Bergquist P.L.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            91818 PRELIMINARY; PRT; 091818; 001007-2000 (TrEMBLrel. 15, Created) 01-007-2000 (TrEMBLrel. 15, Last sequil-DEC-2001 (TrEMBLrel. 19, Last anno
39 QVGCSNVRGSFVKLSTGRTGADYYIE 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        :|| |:| || :: || || 839 QVGCSNLRGSFVKLSTGRTGADYYIE 864
                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  930
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1
34 9
930 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=74586;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 40;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local S
Matches 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                  Q9RFX5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Signal.
                                                                                                                                                                                                                                                     Q9RFX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIGNAL
                                                                                                                                                                7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ٣
```

RESULT 09L8L8

g ð

ä

Gaps

1;

RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT OF RESULT

```
chermophilic anaerobe NA10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HSSP; Q06851; 1NBC
                                                                                                               SEQUENCE FROM N.A.
                                                                  NCBI_TaxID=67756;
                                                                                                                                         STRAIN=NA10;
                                                 Bacteria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09X3P5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              09X3P5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   XXNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
Q9X3P5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DE RAP REP REP DE RAP D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                               ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
STRAIN=TOKTB.1;
MEDLINE-20171169; PubMed-10706665;
Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P.,
Bergquist P.L.;
"Multidomain and multifunctional glycosyl hydrolases from the extreme thermophile Caldicellulosiruptor isolate Tok7B.1.";
Curr. Microbiol. 40:333-340(2000).
                                                                                                                                                                                    846 TGTLEVQYRSGGSGNSSNAVTPQFNLKNTGTQAIDLSTVKIRYYFTKD-GTEELSFWCDY 904
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                            2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                               1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 1751;
                                                                    Length 997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Interior, itrous...,
Pfam: PF00942; CBD_3; 1.
Pfam: PF00792; CBD_3; 1.
PRINTS: PR00759; GIYCO_hydro_9; 1.
PRINTS: PR01217: PRCHEXTENSN.
PRODOM: PD001947; CBD_3; 1.
PROSITE: PS00699; GIXCOSYL, HYDROL_F9_1; 1.
PROSITE: PS00699; GIXCOSYL_HYDROL_F9_2; 1.
PROSITE: PS00213; LIPOCALIN; UNKNOWN.1.
PROSITE: PS00213; LIPOCALIN; UNKNOWN.1.
                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
106927 MW; F20CB119D6410629 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
                                                                Query Match 43.6%; Score 204.5; DB 2; Best Local Similarity 43.7%; Pred. No. 4.6e-14; Matches 38; Conservative 21; Mismatches 27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 41.5%; Score 194.5; DB 2; Best Local Similarity 44.7%; Pred. No. 1.2e-12; Matches 38; Conservative 14; Mismatches 32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CELE.
Caldicellulosiruptor sp. Tok7B.1.
Bacteria; Firmicutes; Bacillus/Clostridium group; Thermoanaerobacter group; Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT; 1000 AA
                                                                                                                                                                                                                                                                                                                                                                                                              PRT; 1751 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HSSP; Q06851; INBC.
InterPro; IPRO01956; CBD_3.
InterPro; IPRO01701; Glyco_hydro_9.
InterPro; IPRO00566; Lipocin_cytrABP.
InterPro; IPRO02965; P_rich_extensn.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last seq
01-DEC-2001 (TrEMBLrel. 19, Last ann
                                                                                                                                                                                                                                                                           | :| |: | :||| |||||:: | 931
                                                                                                                                                                                                                                                 62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              337 IGASNVTFNFVKLSSGVSGADYYLE 761
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
  AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=80339;
266
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              024820
                                                                                                                                                                                                                                                                                                                                                                                                           09A0G4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
024820
                                                                                                                                                                                                                                                                                                                                                                 RESULT
                                                                                                                                                                                                                                                                                                                                                                                      S
                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                 ŏ
```

```
STRAIN=TOK7B.1;
MEDLINE-20171169; PubMed=10706665;
Gibbs M.D. Reeves R.A., Farrington G.K., Anderson P., Williams D.P.,
Bergquist P.L.;
"Multidomain and multifunctional glycosyl hydrolases from the extreme
thermophile caldicellulosiruptor isolate Tok7B.1.";
Curr. Microbiol. 40:333-340(2000).
EMBL; AFOR8737; AAD30363.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
Miyake K., Machida Y., Hattori K., Iijima S.; "Characterization of a multi-domain cellulase from an extremely thermophilic anaerobe strain NA10."; Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases. EMBL, AB008029; BAAA22939.1; -- HSSP; Q06851; INBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 1000;
                                                                                                                                                                                                                                                                                                                        | InterPro; | IPR001956; CBD_3. | |
| InterPro; | IPR001956; CBD_3. |
| InterPro; | IPR001956; CBD_3. |
| InterPro; | IPR001000; Glyco_hydro_10. |
| InterPro; | IPR00150; Glyco_hydro_10. |
| Pfam; | PF00942; CBD_3; 1. |
| Pfam; | PF00942; Cellulase; 1. |
| Pfam; | PF00131; Glyco_hydro_10; 1. |
| Pfam; | PF00134; Glyco_hydro_10; 1. |
| PFINTS; | PR00134; CBD_3; 1. |
| PROSITE; | PS00019; ACTININ.1. | UNKNOWN.1. |
| PROSITE; | PS00591; GLYCOSYL_HYDROL_F10; 1. |
| PROSITE; | PS00591; GLYCOSYL_HYDROL_F5; UNKNOWN.1. |
| PROSITE; | PS00593; GLYCOSYL_F5; UNKNOWN.1. |
| PS00593; GLYCOSYL_F5; UNKNOWN.1. |
| PS00593; GLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   31; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 41.0%; Score 192.5; DB 2; Best Local Similarity 43.7%; Pred. No. 9.9e-13; Matches 38; Conservative 17; Mismatches 31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Caldicellulosiruptor sp. Tok7B.1.
Bacteria: Firmicutes: Bacillus/Clostridium group;
Thermoanaerobacter group; Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | :| |: | |: | | |: 431 AQIGASNVTFKFVKLSSSVSGADYYLE 457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001956; CBD_3.
InterPro; IPR001305; CBD_6.
InterPro; IPR001000; Glyco_hydro_10.
InterPro; IPR002965; P_rich_extensn.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF00942; CBD_3; 3.
Pfam; PF02018; CBD_6; 2.
Pfam; PF00331; Glyco_hydro_10; 1.
PRINTS; PR00134; GLHYDRIASE10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-NOV-1999 (TrEMBLrel. 12, 01-NOV-1999 (TrEMBLrel. 12, 01-DEC-2001 (TrEMBLrel. 19,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
```

ή

4

```
STRAIN=Z-1320;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Zverlov V.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             P96311
P96311;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     P96311
    ò
                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                          pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
                                                                                                                   STRAIN-TOKTB.1;
MEDLINB-20171169; PubMed-10706665;
MEDLINB-20171169; PubMed-10706665;
Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P.,
Bergquist P.L.;
"Multidomain and multifunctional glycosyl hydrolases from the extreme thermophile caldicellulosiruptor isolate Tok7B.1.";
EMBL; AF078040; AAK06391.1;
EMBL; AF078040; AAK06391.1;
InterPro: IPR001956; CBD_3.
Probom; PD001947; CBD_3: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |:|| |||||:: |:| |:| ::|| |||||||| | :||:| ||
118 GLKVLYKNNETSASTGSIRPWFKIVNGGSSSVDLSRVKIRYWYTVDGDKPQSAV-CDWAQ 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                       4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                Length 1770;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    40.8%; Score 191.5; DB 2; Length 261; 43.5%; Pred. No. 2.6e-13;
PRINTS; PR01217; PRICHEXTENSN.
ProDom; PD001947; CBD_3; 3.
PROSITE; PS00591; GLYCOSYL_HYDROL_F10; 1.'
SEQUENCE 1770 AA; 193641 MW; BBAF1937D4926C92 CRC64;
                                                                                   31; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          32; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             261
28759 MW; 4771744A26A6AE04 CRC64;
                                                                                                                                                                                                                                          01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                        GLYCOSYL HYDROLASE 6 (FRAGMENT).
Caldicellulosiruptor sp. Tok78.1.
Bacteria; Firmicutes; Bacillus/Clostridium group;
Thermoanaerobacter group; Caldicellulosiruptor.
                                                              DB 2;
                                                                        Pred. No. 1.9e-12;
                                                                                                                                                                                                                         261 AA.
                                                                                   16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         15; Mismatches
                                                              41.0%; Score 192.5;
43.5%; Pred. No. 1.9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
                                                                                                                                                             1163 IGASNVTFNFVKLSSGVSGADYYLE 1187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     177 IGASNVTFNFVKLSSGVSGADYXLE 201
                                                                                                                                                64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1999 (TrEMBLrel. 12,
01-NOV-1999 (TrEMBLrel. 12,
01-DEC-2001 (TrEMBLrel. 19,
CELB.
                                                                                   37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best_Local Similarity 43.59
Matches 37; Conservative
                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             261 2
261 AA;
                                                                         Best Local Similarity
                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=80339;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         lydrolase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NON_TER
SEOUENCE
                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NON_TER
                                                                                                                                                                                                                       09A0G7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            945X60
                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
Q9X3P6
ID Q9:
                                                                                                                                                                                                             Q9AQG7
  DR DR SO
                                                                                                                           g
                                                                                                                                                                  g
                                                                                                                                                                                                                       ò
                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
1;
                                                                                                                                                                                                                                                   Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P., Bergquist P.L.;
"Multidomain and multifunctional glycosyl hydrolases from the extreme thermophile Caldicellulosiruptor isolate Tok7B.1.";
Curr. Microbiol. 40:333-340(2000).
EMBL, AF078737; AAD30364.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
-1- FUNCTION: THE N-TERMINAL DOMAIN OF CELA ENCODES FOR AN ENDOGLUCANASE ACTIVITY ON CARBOXYMETHYLCELLUIOSE. THE C-TERMINAL DOMAIN PROBABLY ACT SYNERGISTICALLY TO HYDROLYZE CRYSTALLINE CELLULOSE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae; Anaerocellum group; Anaerocellum. NCBI_TaxID=31899;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -:- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC LINKAGES IN CELLULOSE.
-: SIMILARITY: THE N-TERMINAL PART BELONGS TO CELLULASE FAMILY I (FAMILY 9 OF GLYCOSYL HYDROLASES).
-:- SIMILARITY: THE C-TERMINAL PART BELONGS TO CELLULASE FAMILY 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 1426;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PROSITE; PS00019; ACTININ_1; UNKNOWN_1.
PROSITE; PS00591; GIYCOSYL HYDROL_F10;
PROSITE; PS00599; GIYCOSYL HYDROL_F5; UNKNOWN_1.
SEQUENCE 1426 AA; 157544 MW; 29B3FDB85D09A863 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
ENDOGLUCANASE A (FC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE A)
(CELLULASE A) (FRAGMENT).
Caldicellulosiruptor sp. Tok7B.1.
Bacteria: Firmicutes; Bacillus/Clostridium group; Thermoanaerobacter group; Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               43.5%; Pred. No. 1.9e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        40.8%; Score 191.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR001000; Glyco_hydro_10.
InterPro; IPR001547; Glyco_hydro_10.
InterPro; IPR001547; Glyco_hydro_F5.
InterPro; IPR001542; CBD_3; 3.
Pfam; PF00150; cellulase; 1.
Pfam; PF00131; Glyco_hydro_10; 1.
PRINTS; PR00134; GLHYDRLASE10.
PRINTS; PR01217; PRICHEXTENSN.
PRODORING; PR00184; CBD_3; 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001589; Actinin_act_bind.
InterPro; IPR001956; CBD_3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-MAY-1997 (TrEMBLrel. 03, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    472 IGASNVTFNFVKLSSGVSGADYYLE 496
                                                                                                                                                                                                                        MEDLINE-20171169; PubMed-10706665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 40.89
Best Local Similarity 43.59
Matches 37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Anaerocellum thermophilum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q06851; 1NBC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
```

S

```
4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN=Y106; Q1 M., Wang P., Liu X., Qu Y.; Q1 M., Wang P., Liu X., Qu a cellulase from Bacillus subtilis Y106.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-ER69B.1;
Morris D.D., Gibbs M.D., Ford M., Thomas J., Bergquist P.L.;
"Family 10 and 11 xylanase genes from Caldicellulosiruptor sp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 1779;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
40.6%; Score 190.5; DB 2; Length
Best Local Similarity 43.5%; Pred. No. 3.3e-12;
Matches 37; Conservative 14; Mismatches 33; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kylan degradation; Hydrolase; Glycosidase.
SEQUENCE 1779 AA; 194304 MW; CE5269B6806B5CED CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases. EMBL, AF036924; AAB95326.1; -... HSSP, Q06891; INBC.
HIGSP, CORNOTOSE, CBD_3.
InterPro; IPR001905; CBD_6.
InterPro; IPR001000; Glyco_hydro_10.
                                                                                                                                                                                                                           Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                           Bacteria, Firmicutes, Bacillus/Clostridium group,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacillus subtilis.
Bacteria; Firmicutes; Bacillus/Clostridium group;
                                                                                                                                                                                  PRT; 1779 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Problem: Pro0942; CBD_3; 3.
Pfam; Pr00942; CBD_6; 2.
Pfam; Pr00331; Glyco_hydro_10; 1.
Prinrs; Pr00134; GLHYDRLASE10.
Probom; PD001947; CBD_3; 3.
PROSITE; PS00591; GLYCOL_PIOCOL_10; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Bacillus/Staphylococcus group; Bacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           903 IGASNVTFNFVKLSSGVSGADYYLE 927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                        Caldicellulosiruptor sp. Rt69B.1.
                                                                                                                                                                                                                           01-JUN-1998 (TEMBLrel. 06, Creat
01-JUN-1998 (TEMBLrel. 06, Last
01-DEC-2001 (TEMBLrel. 19, Last
PAMILY 10 XYLANASE (EC 3.2.1.8).
                          64 MGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19, ALKALI TOLERABLE CELLULASE.
                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1779 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=70295;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q93TJ6;
                                                                                                                                                                                052374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q93TJ6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Dp
                                                                  g
                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                          δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           .
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "Multidomain and multifunctional glycosyl hydrolases from the extreme thermophile Caldicellulosiruptor isolate Tok7B.1.";
Curr. Microbiol. 40:333-340(2000).
EMBL: AF078038; AAK05388.1; -.
HSSP; Q06851; INBC.
InterPro; IPR002866; BAR.
InterPro; IPR001956; CBD_3.
InterPro; IPR002965; P. rich_extensn.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P., Bergquist P.L.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 VSGG-VKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNC 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ;
;
                                                         InterPro; IPR001956; GBD_3.
InterPro; IPR001556; Glyco_hydro_48.
InterPro; IPR001701; Glyco_hydro_9.
Pfam; PF00101; Glyco_hydro_9.
Pfam; PF00101; Glyco_hydro_9; 1.
Pfam; PF00101; Glyco_hydro_9; 1.
ProDom; PF00194; GLHVDRLASE48.
ProDom; PD011903; Glyco_hydro_18; 1.
ProDom; PD011903; Glyco_hydro_48; 1.
PROSITE; PS00592; GLYCOSYL_HYDROL_F9_1; 1.
PROSITE; PS00698; GLYCOSYL_HYDROL_F9_1; 1.
PROSITE; PS00698; GLYCOSYL_HYDROL_F9_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 996;
                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match

40.7%; Score 191; DB 2; Length 1711;
Best Local Similarity 43.8%; Pred. No. 2.7e-12;
Matches 39; Conservative 18; Mismatches 30; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                              1711 AA; 189979 MW; E3E987CEB9CD0C21 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3C72B6ED22F3C614 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-50N-2001 (TrEMBLrel. 17, Created)
01-50N-2001 (TrEMBLrel. 17, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
GLYCOSYL HYDROLASE 5 (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 40.6%; Score 190.5; DB 2
Best Local Similarity 43.5%; Pred. No. 1.6e-12;
Matches 37; Conservative 15; Mismatches 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Caldicellulosituptor sp. Tok7B.1.
Bacteria; Firmicutes; Bacillus/Clostridium grou
Thermoanaerobacter group; Caldicellulosiruptor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              996 AA.
(FAMILY 48 OF GLYCOSYL HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                742 DWAQIGASNVTFKFVKLSSSVSGADYYLE 770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 60 DWAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-TOK7B.1;
MEDLINE-20171169; Pubmed-10706665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               996
108275 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRINTS; PR01217; PRICHEXTENSN. ProDom; PD001947; CBD_3; 1.
                   EMBL; Z86105; CAB06786.1;
HSSP; P26221; 1TF4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               996 9
996 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=80339;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hydrolase.
                                                                                                                                                                                                                                                                                                                                        NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9AQH0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09AQH0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Dp
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCOORDING TO BE SENT OF SENT 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
```

ij

Gaps

ij

```
q
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                       οy
                                                                                             ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .;
;
                                                                                                                                    Gaps
                                                                                                                      4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-79-23;
Jung K.H., Chun Y.C., Lee J.-C., Kim J.H., Yoon K.-H.;
"Cloning and expression of a Bacillus sp. 79-23 cellulase gene.";
Biotechnol. Lett. 18:1077-1082(1996).
                                                                  Length 499;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        40.5%; Score 190; DB 2; Length 508; 40.7%; Pred. No. 8.3é-13; Live 21; Mismatches 26; Indels
                                                                Query Match 40.5%; Score 190; DB 2; Length 49
Best Local Similarity 40.7%; Pred. No. 8.1e-13;
Matches 35; Conservative 21; Mismatches 26; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF355629; AAK39540.1; -.
SEQUENCE 499 AA; 55252 MW; 587A687584B7E5BE CRC64;
                                                                                                                                                                                                                                                                                  Q93LD0;
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
ENDO-1,4-BETA-GLUCANASE (EC 3.2.1.4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JUN-1998 (TrEMBLrel. 06, Created)
01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
ENDO-B-1,4-GLUCANASE (EC 3.2.14).
                                                                                                                                                                                                                                                                                                                                                               Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
NCBL_TaxID=1423;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Bacillus sp. 79-23.
Bacteria: Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
NCBI_TaxID-72363;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               499 AA.
                                                                                                                                                                                           410 QIGCGNVTHKFVTLHKPKQGADTYLE 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                             63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity 40.7 tes 35; Conservative
                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                     Bacillus subtilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            052731;
                                                                                                                                                                                                                                           RESULT 14
1931D0
AC 0931D0
AC 0931D0
DT 01-DEC
DT 01-DEC
DT 01-DEC
DT 01-DEC
OC Bacill
OC Bacill
OC Bacill
OC Bacill
OC STRUIN
RN (1)
RN SEQUEN
RL SUBmit
RN SUBmit
                                                                                                                                                                                                                                                                      093LD0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              052731
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 15
052731
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
 S Rr
So
                                                                                                                                                 a
                                                                                                                                                                             ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
```

```
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                   4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA 62
                                                                                                                                                                                                                                                                                                                                                                                                           4
                                                                                                                                                                                                                                                                                                                                                              Length 499;
                                     Yoon K.-H., Jung K.H.;
Submitted (JAN.1998) to the EMBL/GenBank/DDBJ databases.
EMBL; AR6045482; AAC02536.1; -.
HSSP; O85465; LA3H.
InterPro; IPR001956; CBD_3.
InterPro; IPR001547; Glyco_hydro_F5.
Pfam; PF00942; CBD_3.
Pfam; PF00150; cellulase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                        55193 MW; 990A1878788CF738 CRC64;
                                                                                                                                                                                                                       Prodom; PD001947; CBD_3; 1.
PROSTIE: PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
HYDRODENGE; Glycosidase.
SEQUENCE 499 AA: 55193 WW; 990A1878788CF738
                                                                                                                                                                                                                                                                                                                                                       Score 187; DB 2;
Pred. No. 1.8e-12;
0; Mismatches 27.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Search completed: July 2, 2002, 09:17:35
Job time: 343 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIGCONVTHKEVTLHKPKQGADTYLE 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                          39.9%; Sco
ilarity 40.7%; Pro
Conservative 20;
                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity
Matches 35; Conserv
SEQUENCE FROM N.A.
```

THIS PAGE BLANK (USPTO)